

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 08:07:49 ; Search time 42.56 Seconds
(without alignments)
11.827 Million cell updates/sec

Title: US-09-580-201A-2
Perfect score: 73
Sequence: 1 NGVCGXXLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	45	61.6	400	1 VG78_HSVII	Q00167 ictaluriid h
2	44	60.3	1235	1 DNBI_HCMVA	P17147 human cytom
3	43	58.9	746	1 CLC5_HUMAN	P51795 homo sapien
4	43	58.9	746	1 CLC5_MOUSE	Q9wvd4 mus musculu
5	43	58.9	746	1 CLC5_RAT	P17936 rattus norv
6	42	57.5	172	1 PGPA_ECOLI	P18200 escherichia
7	41.5	56.8	768	1 ITB8_RABIT	P26013 oryctolagus
8	41	56.2	161	1 NEUV_CHICK	P24787 gallus gall
9	41	56.2	554	1 ARDI_RAT	P36407 rattus norv
10	41	56.2	574	1 ARDI_HUMAN	P36406 homo sapien
11	40	54.8	837	1 UBPA_DICDI	P54201 dictyosteli
12	39	53.4	55	1 M8AC_DROME	Q01644 drosophila
13	39	53.4	144	1 NEU2_CAVPO	P10769 cavia porce
14	39	53.4	245	1 IE0_NPVOP	O10369 orgyia pseu
15	39	53.4	266	1 GNPI_GIALA	O97439 giardia lam
16	39	53.4	1160	1 DNBI_SCMVC	P13215 simian cyto
17	38.5	52.7	3175	1 RPOA_EAV	P19811 equine arte
18	38	52.1	155	1 NEU4_CATCO	P16229 catostomus
19	38	52.1	1191	1 DNBI_MCMVS	P30672 murine cyto
20	37	50.7	96	1 FSPM_LYCES	P14903 lycopersico
21	37	50.7	125	1 NEU1_BOVIN	P01175 bos taurus
22	37	50.7	125	1 NEU1_SHEEP	P13389 ovis aries
23	37	50.7	196	1 V17_BPT7	P03781 bacterioph
24	37	50.7	392	1 FZD9_CHICK	Q91802 gallus gall
25	36.5	50.0	1808	1 TENA_CHICK	P10039 gallus gall
26	36	49.3	49	1 S325_SEGFL	P58605 segestria f
27	36	49.3	88	1 V184_FOPVP	Q9j549 fowlpox vir
28	36	49.3	131	1 NEU2_ANSAN	P19630 anser anser
29	36	49.3	186	1 DHML_METFL	O50425 methylobaci
30	36	49.3	187	1 DHML_METME	Q59543 methylobaci
31	36	49.3	206	1 CLD5_RAT	Q9jko6 rattus norv
32	36	49.3	218	1 CLD5_MOUSE	O54942 mus musculu
33	36	49.3	233	1 TPL_TREPA	P45685 treponema p

34	36	49.3	244	1 CSMI_CHLTE	O68988 chlorobium
35	36	49.3	250	1 YS85_CAEEL	Q09623 caenorhabdi
36	36	49.3	261	1 IE0_NPVAC	P41710 autographa
37	36	49.3	309	1 MAT1_HUMAN	P51948 homo sapien
38	36	49.3	309	1 MAT1_MOUSE	P51949 mus musculu
39	36	49.3	309	1 MAT1_XENLA	P51951 xenopus lae
40	36	49.3	354	1 TFF1_ALCEU	P27137 alcaligenes
41	36	49.3	441	1 GDF9_MOUSE	Q07105 mus musculu
42	36	49.3	520	1 SAT2_HUMAN	Q9um82 homo sapien
43	36	49.3	1127	1 TF1G_HUMAN	Q9upn9 homo sapien
44	36	49.3	3051	1 YNX3_CAEEL	P34576 caenorhabdi
45	35.5	48.6	25	1 CX44_CONPU	P55963 conus purpu

ALIGNMENTS

RESULT 1					
VG78_HSVII					
ID	VG78_HSVII	STANDARD;	PRT;	400 AA.	
AC	Q00167:				
DT	01-DEC-1992 (Rel. 24, Created)				
DT	01-DEC-1992 (Rel. 24, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	Hypothetical gene 78 zinc-binding protein.				
GN	78.				
OS	Ictaluriid herpesvirus 1 (Channel catfish virus) (CCV).				
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;				
OC	unclassified Herpesviridae.				
OX	NCBI_TaxID=10401;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN=AUBURN 1;				
RX	MEDLINE=92087490; PubMed=1727613;				
RA	Davidson A.J.;				
RT	"Channel catfish virus: a new type of herpesvirus.";				
RL	Virology 186:9-14(1992).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; M75136; AAA88180.1; -				
DR	PIR; D36794; ZBBE14.				
KW	Hypothetical protein; zinc; zinc-finger.				
SQ	SEQUENCE 400 AA; 44167 MW; D941DB7738B95CA9 CRC64;				

Query Match 61.6%; Score 45; DB 1; Length 400;
Best Local Similarity 60.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	4	CCGXXLCHXC	13
DB	133	CCGATLCDSC	142

RESULT 2					
DNBI_HCMVA					
ID	DNBI_HCMVA	STANDARD;	PRT;	1235 AA.	
AC	P17147;				
DT	01-AUG-1990 (Rel. 15, Created)				
DT	01-AUG-1990 (Rel. 15, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Major DNA-binding protein (MDBP).				
GN	UL57 OR DBP.				
OS	Human cytomegalovirus (strain AD169).				
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Betaherpesvirinae; Cytomegalovirus.				

OX NCBI_TaxID=10360;
 RP [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
 RA "Analysis of the protein-coding content of the sequence of human
 RA cytomegalovirus strain AD169.";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -1- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
 CC REPLICATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X17403; CAA35372.1; -;
 DR PIR; S09820; Q0BEM4.
 DR InterPro; IPR000635; Viral_DNA_bind.
 DR Pfam; PF00747; Viral_DNA_bp; 1.
 KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein;
 KW Early protein.
 FT ZN_FING 467 481 C2HC-TYPE.
 SQ SEQUENCE 1235 AA; 133878 MW; 94E8D4F8D3BB2CB6 CRC64;

Query Match 60.3%; Score 44; DB 1; Length 1235;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 7; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 2 GVC--CGXXLCHXC 13
 I:I I I I I
 Db 470 GLCEACGGTCGTC 483

RESULT 3
 CLC5_HUMAN STANDARD; PRT; 746 AA.
 AC P51795;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Chloride channel protein 5 (CLC-5).
 GN CLC5 OR CLCK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Kidney;
 RX MEDLINE=96121370; PubMed=8575751;
 RA Fisher S.E., van Bakel I., Lloyd S.E., Pearce S.H.S.,
 RA Thakker R.V., Craig I.W.;
 RT "Cloning and characterization of CLCN5, the human kidney chloride
 RT channel gene implicated in Dent disease (an X-linked hereditary
 RT nephrolithiasis).";
 RL Genomics 29:598-606(1995).
 RN [2]
 RP SEQUENCE OF 487-746 FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=95179126; PubMed=7874126;
 RA Fisher S., Black G.C.M., Lloyd S.E., Hatchwell E., Wrong O.,
 RA Thakker R.V., Craig I.W.;
 RT "Isolation and partial characterization of a chloride channel gene

RT which is expressed in kidney and is a candidate for Dent's disease
 RL (an X-linked hereditary nephrolithiasis).";
 RN Hum. Mol. Genet. 3:2053-2059(1994).
 RP [3]
 RP TISSUE SPECIFICITY.
 RC TISSUE=Vascular smooth muscle, and Aortic endothelium;
 RX MEDLINE=99222497; PubMed=10198195;
 RA Lamb F.S., Clayton G.H., Liu B.-X., Smith R.L., Barna T.J.,
 RA Schutte B.C.;
 RT "Expression of CLCN voltage-gated chloride channel genes in human
 RT blood vessels.";
 RL J. Mol. Cell. Cardiol. 31:657-666(1999).
 RN [4]
 RN VARIANTS NPHL ARG-200; LEU-244; GLU-506 AND PRO-520.
 RX MEDLINE=96158876; PubMed=859248;
 RA Lloyd S.E., Pearce S.H.S., Fisher S.E., Steinmeyer K., Schwappach B.,
 RA Scheinman S.J., Harding B., Bolino A., Devoto M., Goodyer P.,
 RA Rignen S.P.A., Wrong O., Jentsch T.J., Craig I.W., Thakker R.V.;
 RT "A common molecular basis for three inherited kidney stone diseases.";
 RL Nature 379:445-449(1996).
 RN [5]
 RN VARIANTS NPHL.
 RX MEDLINE=97402204; PubMed=9259268;
 RA Lloyd S.E., Guenther W., Pearce S.H.S., Thomson A., Bianchi M.L.,
 RA Bosio M., Craig I.W., Fisher S.E., Scheinman S.J., Wrong O.,
 RA Jentsch T.J., Thakker R.V.;
 RT "Characterisation of renal chloride channel, CLCN5, mutations in
 RT hypercalcaemic nephrolithiasis (kidney stones) disorders.";
 RL Hum. Mol. Genet. 6:1233-1239(1997).
 CC -1- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
 CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
 CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
 CC TRANSEPITHELIAL TRANSPORT. MAY PLAY AN IMPORTANT ROLE IN RENAL
 CC TUBULAR FUNCTION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: KIDNEY. MODERATELY EXPRESSED IN AORTIC
 CC VASCULAR SMOOTH MUSCLE AND ENDOTHELIAL CELLS, AND AT A SLIGHTLY
 CC HIGHER LEVEL IN THE CORONARY VASCULAR SMOOTH MUSCLE.
 CC -1- DISEASE: DEFECTS IN CLCN5 ARE THE CAUSE OF FOUR DISORDERS OF
 CC HEREDITARY HYPERCALCAEMIC NEPHROLITHIASIS (KIDNEY STONES), THAT
 CC HAVE BEEN REFERRED TO AS DENT'S DISEASE (DD), X-LINKED RECESSIVE
 CC NEPHROLITHIASIS (XRN), X-LINKED RECESSIVE HYPOPHOSPHATEMIC
 CC RICKETS (XLRH) AND IDIOPATHIC LOW MOLECULAR WEIGHT PROTEINURIA OF
 CC JAPANESE CHILDREN (JILP). ALL FOUR DISEASES REPRESENT RENAL
 CC TUBULAR DISORDERS THEY ARE CHARACTERIZED BY LOW MOLECULAR WEIGHT
 CC PROTEINURIA, HYPERCALCAEMIA, NEPHROCALCINOSIS, NEPHROLITHIASIS
 CC (KIDNEY STONES) AND RENAL FAILURE. DD IS A FORM OF FANCONI
 CC SYNDROME (ALSO KNOWN AS X-LINKED RECESSIVE NEPHROLITHIASIS TYPE 2
 CC (NPHL2)).
 CC -1- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91906; CAA63000.1; -;
 DR EMBL; X81836; CAA57430.1; -;
 DR MIM; 300008; -;
 DR MIM; 300009; -;
 DR MIM; 310468; -;
 DR InterPro; IPR000644; CBS.
 DR InterPro; IPR001807; Volt_Cl_channel.
 DR Pfam; PF00571; CBS; 2.
 DR Pfam; PF00654; voltage_CLC; 1.
 DR PRINTS; PR00762; CLCHANNEL.
 DR SMART; SM00116; CBS; 2.
 KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
 KW CBS domain; Repeat; Disease mutation.

```

FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 521 541 POTENTIAL.
FT DOMAIN 585 645 CBS 1.
FT DOMAIN 682 733 CBS 2.
FT VARIANT 30 30 /FTId-VAR_001615.
FT VARIANT 57 57 G -> V (IN DD).
FT VARIANT 200 200 /FTId-VAR_001616.
FT VARIANT 244 244 L -> R (IN NPHL2).
FT VARIANT 280 280 /FTId-VAR_001617.
FT VARIANT 506 506 S -> L (IN NPHL3).
FT VARIANT 512 512 R -> P (IN JILP).
FT VARIANT 520 520 /FTId-VAR_001619.
FT VARIANT 527 527 G -> E (IN NPHL1).
FT VARIANT 527 527 /FTId-VAR_001620.
FT VARIANT 520 520 CURRENTS).
FT VARIANT 527 527 /FTId-VAR_001621.
FT VARIANT 527 527 S -> P (IN NPHL2).
FT VARIANT 527 527 /FTId-VAR_001622.
FT VARIANT 527 527 E -> D (IN DD; ABOLISHES THE CHLORIDE
FT VARIANT 527 527 CURRENTS AND TOTAL LOSS OF FUNCTION).
FT VARIANT 527 527 /FTId-VAR_001623.
FT SEQUENCE 746 AA; 83146 MW; EF913C5BA40C85D8 CRC64;

Query Match 58.9%; Score 43; DB 1; Length 746;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
   ||| |||
Db 220 CCGNILCH 227

RESULT 4
CLC5_MOUSE
ID CLC5_MOUSE STANDARD; PRT; 746 AA.
AC Q9WVD4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 01-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Chloride channel protein 5 (CLC-5).
GN CLCN5 OR CLC5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=99303559; PubMed=10373326;
RA Tanaka K., Fisher S.E., Craig I.W.;
RT "Characterization of novel promoter and enhancer elements of the mouse
RT homologue of the Dent disease gene, CLCN5, implicated in X-linked
RT hereditary nephrolithiasis."
RL Genomics 58:281-292(1999).
CC -1- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
CC TRANSEPITHELIAL TRANSPORT. MAY PLAY AN IMPORTANT ROLE IN RENAL
CC TUBULAR FUNCTION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.

```

```

CC -1- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
CC -1- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AF134117; AAD28473.1; -.
DR MGD; MGI:99486; CLCN5.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR001807; Volt_CL_channel.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00634; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW CBS domain; Repeat.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 494 514 POTENTIAL.
FT TRANSMEM 521 541 CBS 1.
FT DOMAIN 585 641 CBS 2.
FT DOMAIN 682 733 CBS 2.
FT SEQUENCE 746 AA; 83100 MW; D8F3AE4FFC331A08 CRC64;

Query Match 58.9%; Score 43; DB 1; Length 746;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
   ||| |||
Db 220 CCGNILCH 227

RESULT 5
CLC5_RAT
ID CLC5_RAT STANDARD; PRT; 746 AA.
AC P51796; P70642;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Chloride channel protein 5 (CLC-5).
GN CLCN5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=96125100; PubMed=8537381;
RA Steimeyer K., Schwappach B., Bens M., Vandewalle A., Jentsch T.J.;
RT "Cloning and functional expression of rat CLC-5, a chloride channel
RT related to kidney disease."
RL J. Biol. Chem. 270:31172-31177(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=96215316; PubMed=8626585;
RA Sakamoto H., Kawasaki M., Uchida S., Sasaki S., Marumo F.;
RT "Identification of a new outwardly rectifying Cl- channel that
RT belongs to a subfamily of the ClC Cl- channels."

```

J. Biol. Chem. 271:10210-10216(1996).

-!- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME; MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND TRANSEPITHELIAL TRANSPORT. MAY PLAY AN IMPORTANT ROLE IN RENAL TUBULAR FUNCTION.

-!- SUBCELLULAR LOCATION: Integral membrane protein.

-!- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.

-!- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.

-!- SIMILARITY: CONTAINS 2 CBS DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z56277; CAA91216.1; -;
 EMBL; D50497; BAA09091.1; -;
 InterPro: IPR000644; CBS.
 InterPro: IPR001807; Volt_CL_channel.
 DR Pfam; PF00571; CBS; 2.
 DR Pfam; PF00654; voltage_CLC; 1.
 DR PRINTS; PR00762; CLCHANNEL.
 DR SMART; SM00116; CBS; 2.
 KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
 CBS domain; Repeat. 75
 FT TRANSMEM 55 POTENTIAL.
 FT TRANSMEM 133 153 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT TRANSMEM 246 266 POTENTIAL.
 FT TRANSMEM 317 337 POTENTIAL.
 FT TRANSMEM 352 372 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT TRANSMEM 453 473 POTENTIAL.
 FT TRANSMEM 494 514 POTENTIAL.
 FT TRANSMEM 521 541 POTENTIAL.
 FT TRANSMEM 585 614 CBS 1.
 FT DOMAIN 682 733 CBS 2.
 FT CONFLICT 315 315 H -> Y (IN REF. 2).
 SQ SEQUENCE 746 AA; 83067 MW; 5F17D45F397003CE CRC64;

Query Match 58.9%; Score 43; DB 1; Length 746;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
 ||| |||
 Db 220 CCGNLTCH 227

RESULT 6
 PGPA_ECOLI
 ID PGPA_ECOLI STANDARD; PRT; 172 AA.
 AC P18200; P77321;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphatidyglycerophosphatase A (EC 3.1.3.27).
 GN PGPA OR B0418.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 ON NCBI_TaxID=562;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89033892; PubMed=2846510;
 RA Icho T.;
 RT "Membrane-bound phosphatases in Escherichia coli: sequence of the pgpA gene.";

J. Bacteriol. 170:5110-5116(1988).

[2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=K12 / MG1655;
 RC MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).

[3]
 RN SEQUENCE FROM N.A.
 RP Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
 RA Duncan M., Federspiel N., Hyman R., Kaiman S., Komp C., Kurdi O.,
 RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
 RA Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

[4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=K12 / W3110;
 RA Iida A., Hayashi M., Fujio T., Teshiba S.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: ONE OF THE THREE PHOSPHOLIPID PHOSPHATASES, SPECIFICALLY HYDROLYZES PHOSPHATIDYGLYCEROPHOSPHATE.

-!- CATALYTIC ACTIVITY: Phosphatidyglycerophosphate + H(2)O = phosphatidylglycerol + phosphate.

-!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

-!- CAUTION: REF. 1 SEQUENCE DIFFERS FROM THAT SHOWN BY A FRAMESHIFT IN POSITION 42. IN ADDITION THE AUTHOR OF REF. 1 HAS TRANSLATED THE WRONG DNA STRAND THUS PRODUCING AN ORF WHICH HAS NOTHING TO DO WITH THE ONE SHOWN HERE.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M23546; AAA24325.1; ALT_SEQ.
 EMBL; AE000148; AAC73521.1; -;
 EMBL; U82664; AAB40174.1; -;
 EMBL; D17333; BAA21779.1; -;
 PIR; A30192; PAECGA.
 DR EcoGene; EG10704; pgpA.
 KW Hydrolase; Phospholipid degradation; Transmembrane; Inner membrane;
 KW Complete proteome.
 FT TRANSMEM 32 52 POTENTIAL.
 FT TRANSMEM 54 74 POTENTIAL.
 FT TRANSMEM 142 162 POTENTIAL.
 SQ SEQUENCE 172 AA; 19418 MW; 9DA1C817CA36C8B9 CRC64;

Query Match 57.5%; Score 42; DB 1; Length 172;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11
 |:| | |||
 Db 67 GICIGVYLCH 76

RESULT 7
 ITB8_RABIT
 ID ITB8_RABIT STANDARD; PRT; 768 AA.
 AC P26013;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Integrin beta-8 precursor.
 GN ITGB8.
 OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEW ZEALAND WHITE; TISSUE=Placenta;
RX MEDLINE=92011767; PubMed=1918072;
RA Moyle M., Napier M.A., McLean J.W.;
RT "Cloning and expression of a divergent integrin subunit beta 8";
RL J. Biol. Chem. 266:19650-19658(1991).
CC -1- FUNCTION: INTEGRIN ALPHA-V/BETA-8 IS A RECEPTOR FOR FIBRONECTIN.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-8
CC ASSOCIATES WITH ALPHA-V.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PLACENTA, KIDNEY, BRAIN, OVARY, UTERUS, AND IN
CC SEVERAL TRANSFORMED CELLS.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 VWFA-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M73781; AAA31280.1; -;
DR PIR; B41029; B41029.
DR HSSP; P04070; 1AUF.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002369; Integrin_B.
DR InterPro; IPR01169; Integrin_beta_C.
DR InterPro; IPR003659; PSI.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF00362; Integrin_B; 1.
DR ProDom; PD00181; Integrin_B; 1.
DR SMART; SM00181; EGF; 1.
DR -SMART; SM00187; INB; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00327; VW; 1.
DR PROSITE; PS00243; INTEGRIN_BETA; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
KW Repeat; Signal.
FT SIGNAL 1 42
FT CHAIN 43 768
FT DOMAIN 43 683
FT TRANSMEM 684 703
FT DOMAIN 704 768
FT DOMAIN 146 384
FT DOMAIN 471 629
FT REPEAT 471 510
FT REPEAT 511 552
FT REPEAT 553 592
FT REPEAT 593 629
FT REPEAT 47 469
FT DISULFID 47 65
FT DISULFID 55 65
FT DISULFID 58 94
FT DISULFID 68 83
FT DISULFID 211 218
FT DISULFID 266 307
FT DISULFID 407 419
FT DISULFID 467 471
FT DISULFID 481 491
FT DISULFID 494 499
FT DISULFID 526 531
FT DISULFID 528 561
FT DISULFID 533 546
FT DISULFID 567 572
FT DISULFID 574 583
FT DISULFID 585 593
BY SIMILARITY.

FT DISULFID 607 612
FT DISULFID 609 657
FT DISULFID 614 624
FT DISULFID 627 630
FT DISULFID 634 643
FT CARBOHYD 233 233
FT CARBOHYD 402 402
FT CARBOHYD 421 421
FT CARBOHYD 431 431
FT CARBOHYD 456 456
FT CARBOHYD 648 648
SQ SEQUENCE 768 AA; 84405 MW; C3F6DF0322FFB544 CRC64;

Query Match 56.8%; Score 41.5; DB 1; Length 768;
Best Local Similarity 72.7%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Oy 2 GVC-CGXXLCH 11
||| || |||
Db 524 GVCVCGKCLCH 534

RESULT 8
NEUV_CHICK
ID NEUV_CHICK STANDARD; PRT; 161 AA.
AC P24787;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Vasotocin-neurophysin VT precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=Hypothalamus;
RA Hunt N., Kluever D., Ivell R.;
RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: VASOTOCIN IS AN ANTI-DIURETIC HORMONE.
CC -1- DOMAIN: IN NON-MAMMALIAN TETRAPODS, THE PROTEOLYTIC PROCESSING OF
CC THE PRO-VASOTOCIN INVOLVES ONLY ONE CLEAVAGE, RELEASING THE
CC HORMONE MOIETY AND A "BIG" NEUROPHYSIN WITH TWO DOMAINS HOMOLOGOUS
CC TO THE MAMMALIAN NEUROPHYSIN II AND COPEPTIN, RESPECTIVELY.
CC -1- PTM: SEVEN DISULFIDE BONDS ARE PRESENT IN NEUROPHYSIN.
CC -1- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X55130; CAA38923.1; -;
DR PIR; S14480; S14480.
DR HSSP; P01180; INPO.
DR InterPro; IPR000981; Neurohypophys_horm.
DR Pfam; PF00220; hormone4; 1.
DR Pfam; PF00184; hormone5; 1.
DR PRINTS; PR00831; NEUROPHYSIN.
DR ProDom; PD001676; Neurohypophys_horm; 1.
DR SMART; SM00003; NH; 1.
DR Hormone; Hypothalamus; Cleavage on pair of basic residues;
KW PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
FT Amidation; Signal.
FT SIGNAL 1 19
FT PEPTIDE 20 28
FT PEPTIDE 32 161
FT DISULFID 20 25
BY SIMILARITY. VASOTOCIN.
BY SIMILARITY. VT NEUROPHYSIN.
BY SIMILARITY.

FT MOD_RES 28 28 AMIDATION (G-29 PROVIDE AMIDE GROUP).
SQ SEQUENCE 161 AA; 16693 MW; 2802FBBED5E52277 CRC64;

Query Match 56.2%; Score 41; DB 1; Length 161;
Best Local Similarity 60.0%; Pred. NO. 14;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NGVCCGXIXC 10
|||
DB 101 NGVCCSADTC 110

RESULT 9
ARDL_RAT
ID ARDL_RAT STANDARD; PRT; 554 AA.
AC P36407;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP-binding protein ARD-1 (Fragment).
GN ARD1 OR ARD-1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93232038; PubMed=8473324;
RA Mishima K., Tsuchiya M., Nighingale M.S., Moss J., Vaughan M.;
RT "ARD 1, a 64-kDa guanine nucleotide-binding protein with a carboxyl-terminal ADP-ribosylation factor domain.";
RL J. Biol. Chem. 268:8801-8807(1993).
CC !- FUNCTION: NOT KNOWN. THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
CC ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT.
CC !- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ARF FAMILY
CC OF GTP-BINDING PROTEINS.
CC !- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC !- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL; L04760; AAA1301.1; -
DR HSSP; P32889; IRRG.
DR InterPro; IPR000251; ARF.
DR InterPro; IPR003649; Bbox_C.
DR InterPro; IPR000315; Znf_bbox.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00025; arf; 1.
DR Pfam; PF00643; zf-B_box; 1.
DR SMART; SM00177; ARF; 1.
DR SMART; SM00502; BBC; 1.
DR SMART; SM00336; BBOX; 2.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01019; ARF; 1.
DR PROSITE; PS00119; ZF_BBOX; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW GTP-binding; zinc-finger.
FT NON_TER 1
FT ZN_FING 11 56 RING-TYPE.
FT ZN_FING 102 148 B BOX-TYPE.
FT DOMAIN 370 554 ARF-LIKE.
FT NP_BIND 391 398 GTP (BY SIMILARITY).
FT NP_BIND 434 438 GTP (BY SIMILARITY).
FT NP_BIND 493 496 GTP (BY SIMILARITY).
SQ SEQUENCE 554 AA; 62187 MW; FB427D6F27680839 CRC64;

Query Match 56.2%; Score 41; DB 1; Length 554;
Best Local Similarity 28.6%; Pred. NO. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

QY 2 GVC-----CGXXLCHXC 13
|||
DB 12 GVCDEVFSIQGKVPRLLLCGHTVCHDC 39

RESULT 10
ARDL_HUMAN
ID ARDL_HUMAN STANDARD; PRT; 574 AA.
AC P36406;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP-binding protein ARD-1.
GN ARD1 OR ARD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93232038; PubMed=8473324;
RA Mishima K., Tsuchiya M., Nighingale M.S., Moss J., Vaughan M.;
RT "ARD 1, a 64-kDa guanine nucleotide-binding protein with a carboxyl-terminal ADP-ribosylation factor domain.";
RL J. Biol. Chem. 268:8801-8807(1993).
CC !- FUNCTION: NOT KNOWN. THE C-TERMINUS CAN ACT AS AN ALLOSTERIC
CC ACTIVATOR OF THE CHOLERA TOXIN CATALYTIC SUBUNIT.
CC !- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ARF FAMILY
CC OF GTP-BINDING PROTEINS.
CC !- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC !- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL; L04510; AAA35940.1; -
DR PIR; A46054; A46054.
DR HSSP; P32889; IRRG.
DR MIN; 601747; -
DR InterPro; IPR000251; ARF.
DR InterPro; IPR003649; Bbox_C.
DR InterPro; IPR000315; Znf_bbox.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00025; arf; 1.
DR Pfam; PF00643; zf-B_box; 1.
DR SMART; SM00177; ARF; 1.
DR SMART; SM00502; BBC; 1.
DR SMART; SM00336; BBOX; 2.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01019; ARF; 1.
DR PROSITE; PS00119; ZF_BBOX; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW GTP-binding; zinc-finger.
FT ZN_FING 31 76 RING-TYPE.
FT ZN_FING 122 168 B BOX-TYPE.
FT DOMAIN 390 574 ARF-LIKE.
FT NP_BIND 411 418 GTP (BY SIMILARITY).
FT NP_BIND 454 458 GTP (BY SIMILARITY).
FT NP_BIND 513 516 GTP (BY SIMILARITY).
SQ SEQUENCE 574 AA; 64066 MW; CB85923B29BF0320 CRC64;

Best Local Similarity 28.6%; Pred. No. 32;
Matches 8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

```

Db      32 GVCEDVFSLQGDKVPRLLLCGHTVCHDC 59
      |||      || :|||

```

CC - I - DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.

Best Local Similarity 54.5%; Pred. No. 57;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

CC MOTIFS.
CC -!- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67703; CA447939.1; -.
DR EMBL; AE003672; AAF54025.1; -.
DR HSSP; P01180; INPO.
DR FlyBase; FBgn0004174; Mst84Dc.
KW Spermatogenesis; Repeat; Multigene family.
SQ SEQUENCE 55 AA; 5225 MW; 95A12F3AEC88BD6C CRC64;

Query Match 53.4%; Score 39; DB 1; Length 55;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13
| | | | |
Db 7 GSCCGYCCGPC 18

RESULT 13
NEU2_CAVPO STANDARD; PRT; 144 AA.
AC P10769;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Vasopressin-neurophysin 2-copeptin precursor [Contains: Arg-
DE vasopressin; Neurophysin 2 (Neurophysin-II); Copeptin].
GN AVP.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RA MEDLINE=87105929; PubMed=3803579;
RX Chauvet M.-T., Rouille Y., Chauvet J., Acher R.;
RT "Guinea pig neurohypophysial hormones. Peculiar processing of the
RT three-domain vasopressin precursor.";
RL FEBS Lett. 210:40-44(1987).
RN [2]
RP SEQUENCE OF 13-144.
RX MEDLINE=87247214; PubMed=3595848;
RA Chauvet J., Chauvet M.-T., Acher R.;
RT "Conformation limited proteolysis in the common neurophysin-copeptin
RT precursor shown by trypsin-sepharose chromatographic proteolysis.";
RL FEBS Lett. 217:180-183(1987).
RN [3]
RP SEQUENCE OF 13-105.
RX MEDLINE=88138574; PubMed=3436704;
RA Chauvet M.-T., Chauvet J., Acher R.;
RT "Guinea pig MSEL-neurophysin. Sequence comparison of eight mammalian
RT MSEL-neurophysins.";
RL Int. J. Pept. Protein Res. 30:676-682(1987).
RN [4]
RP SEQUENCE OF 95-132.
RX MEDLINE=86136563; PubMed=3081370;
RA Chauvet M.-T., Chauvet J., Acher R.;
RT "Guinea pig copeptin. The glycopeptide domain of the vasopressin
RT precursor.";
RL FEBS Lett. 197:169-172(1986).
CC -!- FUNCTION: NEUROPHYSIN 2 SPECIFICALLY BINDS VASOPRESSIN.
CC -!- FUNCTION: VASOPRESSIN HAS A DIRECT ANTIDIURETIC ACTION ON THE
CC KIDNEY, IT ALSO CAUSES VASOCONSTRICTION OF THE PERIPHERAL VESSELS.

DR PIR; A23630; A23630.
DR PIR; A29101; A29101.
DR PIR; S00009; S00009.
DR PIR; JS0300; JS0300.
DR HSSP; P01180; INPO.
DR InterPro; IPR000981; Neurohypophys_horm.
DR Pfam; PF00220; hormones; 1.
DR Pfam; PF00184; hormones; 1.
DR PRINTS; PR00831; NEUROPHYSIN.
DR PRODom; PD001676; Neurohypophys_horm; 1.
DR SMART; SM00003; NH; 1.
DR PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Hypothalamus; Amidation; Cleavage on pair of basic residues;
KW Vasoconstrictor; Glycoprotein.
FT PEPTIDE 1 9 ARG-VASOPRESSIN.
FT PEPTIDE 13 105 NEUROPHYSIN 2.
FT PEPTIDE 107 144 COPEPTIN.
FT DISULFID 1 6
FT DISULFID 22 66 BY SIMILARITY.
FT DISULFID 25 39 BY SIMILARITY.
FT DISULFID 33 56 BY SIMILARITY.
FT DISULFID 40 46 BY SIMILARITY.
FT DISULFID 73 85 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT DISULFID 86 91 BY SIMILARITY.
FT MOD_RES 9 9 AMIDATION (G-10 PROVIDE AMIDE GROUP).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .).
SQ SEQUENCE 144 AA; 15068 MW; CE2B18A162C9ABEA CRC64;

Query Match 53.4%; Score 39; DB 1; Length 144;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NGVCCGXXLC 10
| | | | |
Db 82 NGVCCNDESC 91

RESULT 14
IE0_NPVOP STANDARD; PRT; 245 AA.
AC O10369;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Immediate-early protein IE-0.
GN IE-0.
OS Orgyia pseudotsugata multicapsid polyhedrosis virus (OpMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=164623;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohrmann G.F.;
RT "The sequence of the Orgyia pseudotsugata multinucleocapsid nuclear
RT polyhedrosis virus genome.";
RL Virology 229:381-399(1997).
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U75930; AAC59137.1; -.
DR InterPro; IPR001841; Znf_ring.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.

```
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Early protein; Zinc-finger.
FT ZN_FING 195 240 RING-TYPE.
SQ SEQUENCE 245 AA; 27117 MW; 357F43B15F7B1029 CRC64;

Query Match 53.4%; Score 39; DB 1; Length 245;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 CCGXXLCHXC 13
   |||||
Db 213 CCQYSLCYAC 222

RESULT 15
GNP1_GIALA STANDARD; PRT; 266 AA.
AC 097439;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucosamine-6-phosphate isomerase 1 (EC 3.5.99.6) (Glucosamine-
DE 6-phosphate deaminase 1) (GNPDA 1) (GlcN6P deaminase 1).
GN GPII.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR4 / POLISH GENOTYPE;
RA van Keulen H., Steimle P.A., Bulik D.A., Borowiak R.K., Jarroll E.L.;
RT "Cloning of two Giardia glucosamine 6-phosphate isomerase genes only
RT one of which is transcriptionally activated during encystation.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucosamine 6-phosphate + H(2)O = D-fructose
CC 6-phosphate + NH(3).
CC -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
CC -!- ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF050754; AAD02508.1; -.
DR HSSP; P09375; IDEA.
DR InterPro; IPR000457; Glucosamine_iso.
DR Pfam; PF01182; Glucosamine_iso; 1.
DR PROSITE; PS01161; GLC_GALNAC_ISOMERASE; 1.
KW Hydrolase; Carbohydrate metabolism.
FT ACT_SITE 67 67 BY SIMILARITY.
SQ SEQUENCE 266 AA; 29406 MW; 3EC301B8D868386B CRC64;

Query Match 53.4%; Score 39; DB 1; Length 266;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
   |||
Db 255 CCGSTSCH 262
```

Search completed: May 19, 2002, 08:07:50
Job time: 12216 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 08:06:54 ; Search time 112.09 seconds
(without alignments)
20.064 Million cell updates/sec

Title: US-09-580-201a-2
Perfect score: 73
Sequence: 1 NGVCGGXLCXHC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	63.0	258	12 O36453	O36453 lymantria d
2	46	63.0	514	5 Q95QT3	Q95qt3 caenorhabdi
3	45	61.6	244	12 Q91BQ9	Q91bq9 spodoptera
4	44	60.3	121	11 Q9CZW7	Q9cZW7 mus musculu
5	44	60.3	372	11 Q99ND4	Q99nd4 rattus norv
6	44	60.3	538	4 Q96IF1	Q96if1 homo sapien
7	44	60.3	547	11 P97472	P97472 mus musculu
8	44	60.3	547	11 Q91XC0	Q91xc0 mus musculu
9	43	58.9	104	16 Q9REV9	Q9rev9 pseudomonas
10	43	58.9	310	11 Q9WUG1	Q9wug1 cavia porce
11	43	58.9	563	5 Q9N3Y4	Q9n3Y4 caenorhabdi
12	43	58.9	746	6 Q9TTU3	Q9ttu3 oryctolaqus
13	43	58.9	746	6 Q9GKE7	Q9gke7 sus scrofa
14	43	58.9	746	11 Q99P66	Q99p66 cavia porce
15	43	58.9	808	13 Q9PWK7	Q9pwk7 xenopus lae
16	43	58.9	808	13 O13080	O13080 xenopus lae

17	42	57.5	105	2	O85226	O85226 pseudomonas
18	42	57.5	151	12	O10614	O10614 helicoverpa
19	42	57.5	244	5	O62463	O62463 caenorhabdi
20	42	57.5	244	5	O62464	O62464 caenorhabdi
21	42	57.5	285	12	Q99H38	Q99h38 helicoverp
22	42	57.5	289	12	Q91BK7	Q91bk7 spodoptera
23	42	57.5	322	10	Q39896	Q39896 glycine max
24	42	57.5	326	10	Q39895	Q39895 glycine max
25	42	57.5	332	5	O18012	O18012 caenorhabdi
26	42	57.5	446	4	Q96BQ3	Q96bq3 homo sapien
27	42	57.5	954	4	Q9ULG5	Q9ulg5 homo sapien
28	42	57.5	1166	4	Q9P2R3	Q9p2r3 homo sapien
29	42	57.5	1184	11	O54807	O54807 mus musculu
30	42	57.5	1441	10	Q9LH63	Q9lh63 arabidopsis
31	41	56.2	179	16	Q9HVN1	Q9hvn1 pseudomonas
32	41	56.2	263	16	Q92NF2	Q92nf2 rhizobium m
33	41	56.2	272	5	Q9U1S9	Q9uis9 caenorhabdi
34	41	56.2	328	16	Q97H26	Q97h26 clostridium
35	41	56.2	338	5	Q9U5C2	Q9u5c2 caenorhabdi
36	41	56.2	350	5	Q9N8U4	Q9n8u4 trypanosoma
37	41	56.2	362	5	Q9U483	Q9u483 neospora ca
38	41	56.2	363	5	Q9N5Y3	Q9n5y3 caenorhabdi
39	41	56.2	465	4	O60260	O60260 homo sapien
40	41	56.2	514	10	Q9LD86	Q9ld86 arabidopsis
41	41	56.2	546	4	Q9BZY4	Q9bzy4 homo sapien
42	41	56.2	557	5	Q99654	Q99654 caenorhabdi
43	41	56.2	569	4	Q9BZY5	Q9bzy5 homo sapien
44	41	56.2	721	12	Q91LL5	Q91ll5 white spot
45	40	54.8	98	16	Q9K114	Q9k114 neisseria m

ALIGNMENTS

RESULT 1

O36453 ID O36453 PRELIMINARY; PRT; 258 AA.
AC O36453; Q9YMW3;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE IMMEDIATE EARLY 0 PROTEIN (IMMEDIATE EARLY TRANSACTIVATOR 0).
GN IE-0.
OS Lymantria dispar multicapsid nuclear polyhedrosis virus (LdMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445058; PubMed=9300047;
RA Pearson M.N., Rohrmann G.F.;
RT "Splicing is required for transactivation by the immediate early gene
RT 1 of the Lymantria dispar multinucleocapsid nuclear polyhedrosis
RT virus.";
RL Virology 235:153-165(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99124785; PubMed=9887315;
RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavicek J.M., Rohrmann G.F.;
RT "Sequence and analysis of the genome of a baculovirus pathogenic for
RT Lymantria dispar.";
RL Virology 253:17-34(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavicek J., Rohrmann G.F.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Kuzio J.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006656; AAC58234.1; -.

DR EMBL; AF081810; AAC70206.1; -;
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 258 AA; 29395 MW; AC564CDF92282BAD CRC64;

Query Match 63.0%; Score 46; DB 12; Length 258;
Best Local Similarity 54.5%; Pred. No. 2;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 3 VCCGXLCXHC 13
|||||:|:|
Db 208 VCCGYRVCNAC 218

RESULT 2
Q95QT3 PRELIMINARY; PRT; 514 AA.
AC Q95QT3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 58.6 KDA PROTEIN.
GN C28G1.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Favella A.;
RT "The sequence of C. elegans cosmid C28G1.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41026; AAL02448.1; -;
KW Hypothetical protein.
SQ SEQUENCE 514 AA; 58600 MW; BC33388F0F599446 CRC64;

Query Match 63.0%; Score 46; DB 5; Length 514;
Best Local Similarity 63.6%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 3 VCCGXLCXHC 13
|||||:|:|
Db 263 VTCGHALCHKC 273

RESULT 3
Q9IBQ9 PRELIMINARY; PRT; 244 AA.
AC Q9IBQ9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ORF138 IE0.
OS Spodoptera exigua nucleopolyhedrovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.

OX NCBI_TaxID=10454;
RN [1]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=97437494; PubMed=9292027;
RA van Strien E.A., Faktor O., Hu Z.H., Zuidema D., Goldbach R.W.,
RA Vlask J.M.;
RT "Baculoviruses contain a gene for the large subunit of ribonucleotide
reductase.";
RL J. Gen. Virol. 78:2365-2377(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20036646; PubMed=10567663;
RA IJkel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,
RA Goldbach R.W., Vlask J.M.;
RT "Sequence and organization of the spodoptera exigua multicapsid
nucleopolyhedrovirus genome.";
RL J. Gen. Virol. 80:3289-3304(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Zuidema D.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA IJkel W.F., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
RA Goldbach R.W., Vlask J.M.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF169823; AAF33667.1; -;
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 244 AA; 28719 MW; 1F7662E837A866DB CRC64;

Query Match 61.6%; Score 45; DB 12; Length 244;
Best Local Similarity 60.0%; Pred. No. 2.8;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 4 CCGXXLCXHC 13
|||||:|:|
Db 199 CCGYNLCYVC 208

RESULT 4
Q9CZW7 PRELIMINARY; PRT; 121 AA.
AC Q9CZW7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 10 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2610509F17, FULL INSERT SEQUENCE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.
 DR EMBL; AK012077; BAB28013.1; -.
 DR HSSP; P32965; ICTL.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 2.
 DR ProDom; PD000094; LIM; 2.
 DR SMART; SM00132; LIM; 2.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 2.
 KW LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 121 AA; 13774 MW; 5DB20785DE5F74DE CRC64;

Query Match 60.3%; Score 44; DB 11; Length 121;
 Best Local Similarity 50.0%; Pred. No. 2.3;
 Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC----GXILCHXC 13
 | | | | | | | | |
 DB 91 GCCCFPLDGHLLCHGC 106

RESULT 5
 Q99ND4 PRELIMINARY; PRT; 372 AA.
 ID Q99ND4
 AC Q99ND4; 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE AJUBA PROTEIN (FRAGMENT).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Marie H., Bedford F.K., Moss S.J., Attwell D.;
 RT "The amino terminus of the glial glutamate transporter GLT-1 interacts
 RT with the LIM protein Ajuba.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Marie H.;
 RL Thesis (2000), University College London, University of London,
 RL London, United Kingdom.
 CC -1- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.

DR EMBL; AJ306292; CAC28536.1; -.
 DR HSSP; Q05158; IOLI.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 3.
 DR ProDom; PD000094; LIM; 3.
 DR SMART; SM00132; LIM; 3.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 3.
 KW LIM domain; Metal-binding; Zinc.
 FT NON_TER 1
 SQ SEQUENCE 372 AA; 40210 MW; A6F109D84F41AFCL CRC64;

Query Match 60.3%; Score 44; DB 11; Length 372;
 Best Local Similarity 50.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC----GXILCHXC 13
 | | | | | | | | |
 DB 342 GCCCFPLDGHLLCHGC 357

RESULT 6
 Q96IF1 PRELIMINARY; PRT; 538 AA.
 ID Q96IF1
 AC Q96IF1; 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE SIMILAR TO AJUBA.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC007580; AAH07580.1; -.
 SQ SEQUENCE 538 AA; 56933 MW; 1E0DFA0336976A3C CRC64;

Query Match 60.3%; Score 44; DB 4; Length 538;
 Best Local Similarity 50.0%; Pred. No. 7.9;
 Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC----GXILCHXC 13
 | | | | | | | | |
 DB 508 GCCCFPLDGHLLCHGC 523

RESULT 7
 P97472 PRELIMINARY; PRT; 547 AA.
 ID P97472
 AC P97472; 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE AJUBA.
 GN JUB OR AJUBA.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99263009; PubMed=10330178;
 RA Goyal R.K., Lin P., Kanungo J., Payne A.S., Muslin A.J.,
 RA Longmore G.D.;
 RT "Ajuba, a novel LIM protein, interacts with Grb2, augments mitogen-
 RT activated protein kinase activity in fibroblasts, and promotes meiotic
 RT maturation of Xenopus oocytes in a Grb2- and Ras-dependent manner.";
 RL Mol. Cell. Biol. 19:4379-4389(1999).
 CC -1- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
 CC IONS.

DR EMBL; U79776; AAB38287.1; -.
 DR HSSP; Q05158; IOLI.
 DR MGD; MGI:1341886; Jub.
 DR InterPro; IPR001781; LIM.
 DR Pfam; PF00412; LIM; 3.
 DR ProDom; PD000094; LIM; 3.
 DR SMART; SM00132; LIM; 3.
 DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
 DR PROSITE; PS00023; LIM_DOMAIN_2; 3.
 KW LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 547 AA; 57901 MW; C326772AC1C441B2 CRC64;

Query Match 60.3%; Score 44; DB 11; Length 547;
 Best Local Similarity 50.0%; Pred. No. 8;
 Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC----GXILCHXC 13

Db 517 GCCCFPLDGHLLCHGC 532

RESULT 8

Q91XC0 PRELIMINARY; PRT; 547 AA.
AC Q91XC0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SIMILAR TO AJUBA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SALIVARY GLAND;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC011035; AAH11035.1; -- C326772AC1C155B2 CRC64;
SQ SEQUENCE 547 AA; 57919 MW; C326772AC1C155B2 CRC64;

Query Match 60.3%; Score 44; DB 11; Length 547;

Best Local Similarity 50.0%; Pred. No. 8;

Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC----GXIXLCHXC 13

^ ^ | | | | | | |

Db 517 GCCCFPLDGHLLCHGC 532

RESULT 9

Q9REV9 PRELIMINARY; PRT; 104 AA.
AC Q9REV9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HCNA (HYDROGEN CYANIDE SYNTHASE HCNA).
GN HCNA OR PA2193.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RA Pessi G., Haas D.;
RA "Analysis of the Pseudomonas aeruginosa hydrogen cyanide synthase (hcnABC) regulatory region."
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagie W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen."
RT Nature 406:959-964(2000).
RL EMBL: AF208523; AAF21028.1; --
DR EMBL; AE004646; AAG05581.1; --
DR InterPro: IPR001041; Ferredoxin.
DR Pfam: PF00111; fer2; 1
KW Complete proteome; Iron-sulfur. BB6400658514682 CRC64;
SQ SEQUENCE 104 AA; 11427 MW; BB6400658514682 CRC64;

Query Match 58.9%; Score 43; DB 16; Length 104;

Best Local Similarity 46.2%; Pred. No. 2.9;

Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13

: | | | | | |

Db 56 SGAFCGMGVCHCC 68

RESULT 10

Q9WUG1 PRELIMINARY; PRT; 310 AA.
AC Q9WUG1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CHLORIDE CHANNEL CLC-5 (FRAGMENT).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SMALL INTESTINE EPITHELIUM;
RA Cid L.P., Salinas O., Sepulveda F.V.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF133215; AAD33600.1; --
DR InterPro: IPR001807; Volt_CL_channel.
DR Pfam: PF00654; voltage_CLC; 1.
DR PRINTS: PR00762; CLCHANNEL.
FT NON_TER 1
FT NON_TER 310
SQ SEQUENCE 310 AA; 34691 MW; 15419E709DB531EF CRC64;

Query Match 58.9%; Score 43; DB 11; Length 310;

Best Local Similarity 75.0%; Pred. No. 7.3;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11

| | | | |

Db 165 CCGNILCH 172

RESULT 11

Q9N3Y4 PRELIMINARY; PRT; 563 AA.
AC Q9N3Y4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 63.7 KDA PROTEIN.
GN Y42H9AR.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Tin-Wollam A., Wohldmann P.;
RT "The sequence of C. elegans cosmid Y42H9AR.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024777; AAF0562.3; -;
DR InterPro; IPR000822; Znf-C2H2.
DR InterPro; IPR000306; Znf_FYVE.
DR Pfam; PF01363; FYVE; 1.
DR SMART; SM00064; FYVE; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR DNA-binding; Hypothetical protein; Metal-binding; Zinc-finger.
SQ SEQUENCE 563 AA; 63682 MW; 5864FEBD91432DD3 CRC64;

Query Match 58.9%; Score 43; DB 5; Length 563;
Best Local Similarity 66.7%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 5 CGXXLCHXC 13

Db 204 CGRVLCHSC 212

RESULT 12
O9TTU3
ID O9TTU3 PRELIMINARY; PRT; 746 AA.
AC O9TTU3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE -CHLORIDE CHANNEL CLC-5.
GN CLCN5.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NEW ZEALAND WHITE; TISSUE-CORNEA EPITHELIA;
RA Rae J.L.;
RT "Ion Channels in Cornea Epithelia.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF195523; AAF06018.1; -;
DR InterPro; IPR000644; CBS.
DR InterPro; IPR001807; Volt_CL_channel.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
SQ SEQUENCE 746 AA; 83206 MW; 667F2701C0BF006A CRC64;

Query Match 58.9%; Score 43; DB 6; Length 746;
Best Local Similarity 75.0%; Pred. No. 15;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 CGXXLCH 11

Db 220 CCGNLTCH 227

RESULT 13
O9GKE7
ID O9GKE7 PRELIMINARY; PRT; 746 AA.
AC O9GKE7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OUTWARDLY RECTIFYING CHLORIDE CHANNEL.
GN CLC-5.

OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20545523; PubMed=10978325;
RA Dowland L.K.; Luyckx V.A.; Enck A.H.; Leclercq B.; Yu A.S.L.;
RT "Molecular Cloning and Characterization of an Intracellular Chloride
Channel in the Proximal Tubule Cell Line, LLC-PK1.";
RL J. Biol. Chem. 275:37765-37773(2000).
DR EMBL; AF274055; AAG29104.1; -;
DR InterPro; IPR000644; CBS.
DR InterPro; IPR001807; Volt_CL_channel.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
SQ SEQUENCE 746 AA; 83123 MW; 655E3ED45FC61229 CRC64;

Query Match 58.9%; Score 43; DB 6; Length 746;
Best Local Similarity 75.0%; Pred. No. 15;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 CGXXLCH 11

Db 220 CCGNLTCH 227

RESULT 14
O99P66
ID O99P66 PRELIMINARY; PRT; 746 AA.
AC O99P66;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE CHLORIDE CHANNEL CLCN5.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Cornejo I.; Cid L.P.; Sepulveda F.V.;
RT "Cloning and intestinal expression of guinea pig CLC-5 chloride
channel.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326968; AAG49590.1; -;
DR InterPro; IPR000644; CBS.
DR InterPro; IPR001807; Volt_CL_channel.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
SQ SEQUENCE 746 AA; 83093 MW; 92FC8AFDC7D8D4A5 CRC64;

Query Match 58.9%; Score 43; DB 11; Length 746;
Best Local Similarity 75.0%; Pred. No. 15;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 CGXXLCH 11

Db 220 CCGNLTCH 227

RESULT 15
O9PWK7
ID O9PWK7 PRELIMINARY; PRT; 808 AA.
AC O9PWK7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE CHLORIDE CHANNEL CLC-5.
GN CLC-5.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Mo L., Hellmich H.L., Fong P., Wood T.G., Embesi J., Willis N.K.;
RT "Comparison of amphibian and human CLC-5: Similarity of functional
properties and inhibition by external pH.";
RL J. Membr. Biol. 186:0-0(1999).
DR EMBL; AF063904; AAD24497.1; -.
DR InterPro; IPR000644; CBS; 2.
DR InterPro; IPR001130; TatD.
DR InterPro; IPR001807; Volt_CL_channel.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
DR PROSITE; PS01137; TATD.1; UNKNOWN.1.
SQ SEQUENCE 808 AA; 90068 MW; C3F1ADC759BD1676 CRC64;

Query Match 58.9%; Score 43; DB 13; Length 808;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
||| |||
Db 282 CCGNLLCH 289

Search completed: May 19, 2002, 08:06:55
Job time: 12216 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 04:48:46 ; Search time 60.36 Seconds
(without alignments)
20.695 Million cell updates/sec

Title: US-09-580-201A-2

Perfect score: 73
Sequence: 1 NGVCGXXLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	46	63.0	258	T30368	probable immediate
2	46	63.0	1322	T15689	hypothetical prote
3	45	61.6	400	ZBBE14	44.1K zinc-binding
4	44	60.3	1235	1 QQBEW4	DNA-binding protel
5	43	58.9	104	2 F83370	hydrogen cyanide s
6	42	57.5	171	2 G90687	phosphatidylglycer
7	42	57.5	171	2 C85538	phosphatidylglycer
8	42	57.5	171	2 AD0554	phosphatidylglycer
9	42	57.5	172	2 B64771	phosphatidylglycer
10	42	57.5	244	2 T26913	hypothetical prote
11	42	57.5	244	2 T26912	hypothetical prote
12	42	57.5	322	2 T08592	TCACG-motif-bindin
13	42	57.5	326	2 T08591	TCACG-motif bindin
14	42	57.5	332	2 T24312	hypothetical prote
15	42	57.5	1184	2 T00253	gene Ankhzn protei
16	41.5	56.8	768	2 B41029	integrin beta-8 ch
17	41	56.2	161	2 S14480	arginine-vasotocin
18	41	56.2	179	2 A83068	hypothetical protei
19	41	56.2	328	2 E97129	probable flavodoxi
20	41	56.2	514	2 T10559	hypothetical prote
21	41	56.2	557	2 T27752	hypothetical prote
22	41	56.2	574	2 A46054	GTP-binding protel
23	40	54.8	98	2 E81205	hypothetical prote
24	40	54.8	122	2 D90867	hypothetical prote
25	40	54.8	122	2 E85751	hypothetical prote
26	40	54.8	550	2 G85436	hypothetical prote
27	40	54.8	1101	2 T16840	hypothetical prote
28	39	53.4	55	2 S25774	testis-specific pr
29	39	53.4	144	2 A29101	vasopressin / neur

30	39	53.4	149	2	T21037	hypothetical prote
31	39	53.4	186	2	A45910	ultra-high-sulfur
32	39	53.4	211	2	T20590	hypothetical prote
33	39	53.4	245	2	T10407	immediate early pr
34	39	53.4	286	2	S61199	hypothetical prote
35	39	53.4	303	2	A87616	hydrolase, carbon-
36	39	53.4	1160	2	A36256	DNA-binding protei
37	38.5	52.7	3175	1	RRWVEV	genome polyprotein
38	38	52.1	91	2	S02769	gag 75K protein pr
39	38	52.1	98	2	JC5147	tachycitin precurs
40	38	52.1	155	2	B32669	vasotocin 2 / neur
41	38	52.1	250	2	T29344	hypothetical prote
42	38	52.1	351	2	S76925	hypothetical prote
43	38	52.1	553	2	B88072	protein zK1240.2 l
44	38	52.1	658	2	T33568	hypothetical prote
45	38	52.1	948	2	A57640	retinoblastoma bin

ALIGNMENTS

RESULT 1

T30368

probable immediate-early transactivator 0 - Lymantria dispar nuclear polyhedrosis vi.
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T30368

R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.;

Virology 253, 17-34, 1999

A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantr

A:Reference number: Z20836; MUID:99124785

A:Accession: T30368

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-258 <R0Z>

A:Cross-references: EMBL:AF081810; PIDN:AAC70206.1

C:Keywords: immediate-early protein

Query Match 63.0% Score 46; DB 2; Length 258;
Best Local Similarity 54.5%; Pred. No. 10;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 VCCGXXLCHXC 13

Db 208 VCCGYRVCNAC 218

RESULT 2

T15689

hypothetical protein C28G1.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000

C:Accession: T15689

R:Favell, T.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C28G1.

A:Reference number: Z18389

A:Accession: T15689

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1322 <FAV>

A:Cross-references: EMBL:U41026; NID:g1086701; PID:g1086702; PIDN:AAA82350.1; CESP:C

C:Genetics:

A:Gene: CESP:C28G1.3

A:Introns: 25/1; 131/3; 150/1; 166/3; 180/3; 204/2; 235/3; 344/3; 385/2; 436/3; 483/

C:Superfamily: RING finger homology

F:810-862/Domain: RING finger homology <RRN>

Query Match 63.0% Score 46; DB 2; Length 1322;

Best Local Similarity 63.6%; Pred. No. 31;

Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 VCCGXXLCHXC 13
| | | | |
Db 1071 VTCGHALCHKC 1081

RESULT 3
ZBBE14
44.1K zinc-binding protein - ictalurid herpesvirus 1 (strain auburn 1)
C:Species: ictalurid herpesvirus 1
A:Title: host ictalurus punctatus (channel catfish)
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C:Accession: D36794
R:Davidson, A.J.
submitted to GenBank, January 1992
A:Description: Channel catfish virus: a new type of herpesvirus.
A:Reference number: A36804
A:Accession: D36794
A:Molecule type: DNA
A:Residues: 1-400 <DAV>
R:Davidson, A.J.
Virology 186, 9-14, 1992
A:Title: Channel catfish virus: a new type of herpesvirus.
A:Reference number: A39447; MUID:92087490
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 78
C:Superfamily: ictalurid herpesvirus 44.1K zinc binding protein
C:Keywords: zinc finger

Query Match 61.6%; Score 45; DB 1; Length 400;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 CCGXXLCHXC 13
| | | | |
Db 133 CCGATLGDSC 142

RESULT 4
QOBEW4
DNA-binding protein - human cytomegalovirus (strain AD169)
N:Alternate names: UL57 protein
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C:Accession: S09820
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039
A:Accession: S09820
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1235 <CHE>
A:Cross-references: EMBL:X17403; MID:g59591; PIDN:CAA35372.1; PID:g1780835
A:Note: possible protein-coding frames are given
A:Note: the DNA sequence was submitted to EMBL, December 1989, in computer-readable form
C:Superfamily: herpesvirus DNA-binding protein
C:Keywords: DNA binding

Query Match 60.3%; Score 44; DB 1; Length 1235;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 7; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 2 GVC-CCGXXLCHXC 13
| | | | |
Db 470 GLCEACGGTCCHYC 483

RESULT 5
F83370

hydrogen cyanide synthase HcnA PA2193 [Imported] - Pseudomonas aeruginosa (strain PAO)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83370
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Adaman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337
A:Accession: F83370
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-104 <STO>
A:Cross-references: GB:AE004646; GB:AE004091; NID:g9948213; PIDN:AG05581.1; GSPDB:GN00154
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: hcnA; PA2193

Query Match 58.9%; Score 43; DB 2; Length 104;
Best Local Similarity 46.2%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
| | | | |
Db 56 SGAFCGMGVCHCC 68

RESULT 6
G90687
Phosphatidylglycerophosphatase [Imported] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: G90687
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.; gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G90687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-171 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033894.1; PID:g13359928; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs0471
C:Superfamily: conserved hypothetical protein H11306

Query Match 57.5%; Score 42; DB 2; Length 171;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11
| | | | |
Db 67 GICIGVYLCH 76

RESULT 7
C85538
Phosphatidylglycerophosphatase [Imported] - Escherichia coli (strain O157:H7, substra
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: C85538
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoumis, K.; Apoda
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: C85538

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-171 <STO>

A:Cross-references: GB:AE005174; NID:gl2513273; PIDN:AA054767.1; GSPDB:GN00145; UWGP:Z05

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: pcpA

C:Superfamily: conserved hypothetical protein H11306

Query Match 57.5%; Score 42; DB 2; Length 171;

Best Local Similarity 60.0%; Pred. No. 29;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11

I:| | | | |

Db 67 GICIGVYLCH 76

RESULT 8

AD0554 phosphatidylglycerophosphatase A [imported] - Salmonella enterica subsp. enterica serovar

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AD0554

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AD0554

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-171 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD08876.1; PID:gl6501688; GSPDB:GN00176

C:Genetics:

A:Gene: STY0459

C:Superfamily: conserved hypothetical protein H11306

Query Match 57.5%; Score 42; DB 2; Length 171;

Best Local Similarity 60.0%; Pred. No. 29;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11

I:| | | | |

Db 67 GICIGVYLCH 76

RESULT 9

B64771

phosphatidylglycerophosphatase (EC 3.1.3.27) - Escherichia coli

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 29-Sep-1999

C:Accession: B64771

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A6720; MUID:97426617

A:Accession: B64771

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-172 <BLAT>

A:Cross-references: GB:AE000148; GB:U00096; NID:gl1786614; PIDN:AA073521.1; PID:gl1786620;

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: pcpA

C:Function:

A:Pathway: phospholipid degradation

C:Superfamily: conserved hypothetical protein H11306

C:Keywords: phosphoric monoester hydrolase; transmembrane protein

F:59-75/Domain: transmembrane #status predicted <TM1>

F:143-159/Domain: transmembrane #status predicted <TM2>

Query Match

57.5%; Score 42; DB 2; Length 172;

Best Local Similarity 60.0%; Pred. No. 29;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCH 11

I:| | | | |

Db 67 GICIGVYLCH 76

RESULT 10

T26913

hypothetical protein Y45F10B.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26913

R:McMurray, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20286

A:Accession: T26913

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-244 <WIL>

A:Cross-references: EMBL:AL021487; PIDN:CAA16351.1; GSPDB:GN00022; CESP:Y45F10B.8

A:Experimental source: clone Y45F10B

C:Genetics:

A:Gene: CESP:Y45F10B.8

A:Map position: 4

A:Introns: 90/2; 194/1

Query Match

57.5%; Score 42; DB 2; Length 244;

Best Local Similarity 55.6%; Pred. No. 37;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 CGXXLCHXC 13

I:| | | | |

Db 195 CGHTICHTC 203

RESULT 11

T26912

hypothetical protein Y45F10B.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26912

R:McMurray, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20286

A:Accession: T26912

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-244 <WIL>

A:Cross-references: EMBL:AL021487; PIDN:CAA16350.1; GSPDB:GN00022; CESP:Y45F10B.9

A:Experimental source: clone Y45F10B

C:Genetics:

A:Gene: CESP:Y45F10B.9

A:Map position: 4

A:Introns: 90/2; 194/1

Query Match

57.5%; Score 42; DB 2; Length 244;

Best Local Similarity 55.6%; Pred. No. 37;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13
|| ||| |
Db 195 CGHTICHTC 203

RESULT 12

TGACG-motif-binding protein STF2 - soybean
C:Species: Glycine max (soybean)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08592
R:Cheong, Y.H.; Yoo, C.M.; Park, J.M.; Ryu, G.R.; Goekjian, V.H.; Nagao, R.T.; Key, J.L.
submitted to the EMBL Data Library, September 1995
A:Description: STF1 is a novel TGACG-binding factor with a zinc-finger motif and a bZIP
A:Reference number: Z16445
A:Accession: T08592
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-322 <CHE>
A:Cross-references: EMBL:L28004; NID:g986966; PID:g2934885
A:Experimental source: strain Williams; hypocotyl
C:Genetics:
A:Gene: STF2

Query Match 57.5%; Score 42; DB 2; Length 322;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13
|| ||| |
Db 61 CGFPLCHSC 69

RESULT 13

T08591
TGACG-motif binding protein STF1 - soybean
C:Species: Glycine max (soybean)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08591
R:Cheong, Y.H.; Yoo, C.M.; Park, J.M.; Ryu, G.R.; Goekjian, V.H.; Nagao, R.T.; Key, J.L.
submitted to the EMBL Data Library, September 1995
A:Description: STF1 is a novel TGACG-binding factor with a zinc-finger motif and a bZIP
A:Reference number: Z16445
A:Accession: T08591
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-326 <CHE>
A:Cross-references: EMBL:L28003; NID:g2934883; PID:g2934884
A:Experimental source: strain Williams; hypocotyl

Query Match 57.5%; Score 42; DB 2; Length 326;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13
|| ||| |
Db 61 CGFPLCHSC 69

RESULT 14

T24312
hypothetical protein T01G5.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24312
R:Basham, V.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19873
A:Accession: T24312
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-332 <WIL>
A:Cross-references: EMBL:Z81111; PIDN:CAB03269.1; GSPDB:GN00023; CESP:T01G5.7
A:Experimental source: clone T01G5
C:Genetics:
A:Gene: CESP:T01G5.7
A:Map position: 5
A:Introns: 208/1; 257/1

Query Match 57.5%; Score 42; DB 2; Length 332;
Best Local Similarity 55.6%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13
|| ||| |
Db 209 CGHTMCHTC 217

RESULT 15

T00253
gene Ankhzn protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 05-Nov-1999
C:Accession: T00253
R:Ito, K.; Miyashita, A.; Ishii, N.; Kuriyama, H.; Tomimaga, K.; Saitoh, H.; Maruyama
submitted to the EMBL Data Library, February 1998
A:Description: CDNA cloning of mouse Ankhzn.
A:Reference number: Z14132
A:Accession: T00253
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1184 <ITO>
A:Cross-references: EMBL:AB011370; NID:d1179960; PIDN:BAA24980.1; PID:d1025905
C:Genetics:
A:Note: Ankhzn

Query Match 57.5%; Score 42; DB 2; Length 1184;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 CGXXLCHXC 13
|| ||| |
Db 1144 CGRLCHXC 1152

Search completed: May 19, 2002, 04:48:48
Job time: 10550 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 08:04:48 ; Search time 37.95 seconds
(without alignments)
9.054 Million cell updates/sec

Title: US-09-580-201a-2

Perfect score: 73

Sequence: 1 NGVCCGXXLCHXC 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 107366 seqs, 26430961 residues

Total number of hits satisfying chosen parameters: 107366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*

1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*

2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*

3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*

4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*

5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*

6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*

7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	91.8	13	5	US-09-787-986A-1
2	66	90.4	13	5	US-09-787-986A-3
3	66	90.4	61	5	US-09-787-986A-9
4	61	83.6	13	5	US-09-787-986A-2
5	45	61.6	2119	6	US-10-123-155-61
6	45	61.6	3567	6	US-10-123-155-77
7	44	60.3	654	6	US-10-123-155-441
8	44	60.3	1395	6	US-10-123-155-3
9	44	60.3	1606	6	US-10-123-155-65
10	44	60.3	1606	6	US-10-123-155-257
11	44	60.3	1665	6	US-10-063-545-85
12	44	60.3	1665	6	US-10-063-685-85
13	44	60.3	2276	6	US-10-123-155-9
14	44	60.3	2714	6	US-10-123-155-79
15	44	60.3	3819	6	US-10-123-155-405
16	43	58.9	754	6	US-10-063-545-141
17	43	58.9	754	6	US-10-063-685-141
18	43	58.9	1121	6	US-10-123-155-23
19	43	58.9	1617	6	US-10-063-545-61
20	43	58.9	1617	6	US-10-063-685-61
21	43	58.9	1840	6	US-10-123-155-131
22	43	58.9	2027	6	US-10-123-155-175
23	43	58.9	2159	6	US-10-123-155-87
24	43	58.9	2334	6	US-10-123-155-129
25	43	58.9	2395	6	US-10-063-545-139
26	43	58.9	2395	6	US-10-063-685-139

27	43	58.9	2623	6	US-10-123-155-63
28	43	58.9	2846	6	US-10-063-545-37
29	43	58.9	2846	6	US-10-063-685-37
30	43	58.9	2956	6	US-10-123-155-15
31	43	58.9	4374	6	US-10-123-155-125
32	42	57.5	87	6	US-10-119-235-7
33	42	57.5	635	1	PCT-US02-09944-530
34	42	57.5	1174	6	US-10-123-155-271
35	42	57.5	1395	6	US-10-063-545-9
36	42	57.5	1395	6	US-10-063-685-9
37	42	57.5	1484	6	US-10-123-155-447
38	42	57.5	1570	6	US-10-123-155-455
39	42	57.5	1624	6	US-10-123-155-181
40	42	57.5	1670	6	US-10-123-155-325
41	42	57.5	1869	6	US-10-123-155-269
42	42	57.5	1985	6	US-10-123-155-143
43	42	57.5	2586	6	US-10-063-545-7
44	42	57.5	2586	6	US-10-063-685-7
45	42	57.5	2694	5	US-09-994-404-77

ALIGNMENTS

RESULT 1

US-09-787-986A-1

; Sequence 1, Application US/09787986A

; GENERAL INFORMATION:

; APPLICANT: Lewis, Richard J.

; APPLICANT: Alewood, Paul F.

; TITLE OF INVENTION: NOVEL PEPTIDES

; FILE REFERENCE: Davies Collison Cave

; CURRENT APPLICATION NUMBER: US/09/787,986A

; CURRENT FILING DATE: 2001-03-23

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Conus marmoreus

; FEATURE:

; NAME/KEY: PEPTIDE

; LOCATION: (12)

; OTHER INFORMATION: Xaa at position 12 is 4-hydroxy proline

US-09-787-986A-1

Query Match 91.8%; Score 67; DB 5; Length 13;
Best Local Similarity 84.6%; Pred. No. 0.00082;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
|||||
Db 1 NGVCCGYKLCHXC 13

RESULT 2

US-09-787-986A-3

; Sequence 3, Application US/09787986A

; GENERAL INFORMATION:

; APPLICANT: Lewis, Richard J.

; APPLICANT: Alewood, Paul F.

; TITLE OF INVENTION: NOVEL PEPTIDES

; FILE REFERENCE: Davies Collison Cave

; CURRENT APPLICATION NUMBER: US/09/787,986A

; CURRENT FILING DATE: 2001-03-23

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 13

; TYPE: PRT

; ORGANISM: Conus marmoreus
US-09-787-986A-3

Query Match 90.4%; Score 66; DB 5; Length 13;
Best Local Similarity 76.9%; Pred. No. 0.0011;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NGVCCGXXLCHXC 13
||||| ||| |
Db 1 NGVCCGYKLCHPC 13

RESULT 3

US-09-787-986A-9
; Sequence 9, Application US/09787986A
; GENERAL INFORMATION:
; APPLICANT: Lewis, Richard J.
; APPLICANT: Alewood, Paul F.
; APPLICANT: Sharpe, Iain A.
; TITLE OF INVENTION: NOVEL PEPTIDES
; FILE REFERENCE: Davies Collison Cave
; CURRENT APPLICATION NUMBER: US/09/787,986A
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Conus marmoreus
US-09-787-986A-9

Query Match 90.4%; Score 66; DB 5; Length 61;
Best Local Similarity 76.9%; Pred. No. 0.0041;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NGVCCGXXLCHXC 13
||||| ||| |
Db 49 NGVCCGYKLCHPC 61

RESULT 4

US-09-787-986A-2
; Sequence 2, Application US/09787986A
; GENERAL INFORMATION:
; APPLICANT: Lewis, Richard J.
; APPLICANT: Alewood, Paul F.
; APPLICANT: Sharpe, Iain A.
; TITLE OF INVENTION: NOVEL PEPTIDES
; FILE REFERENCE: Davies Collison Cave
; CURRENT APPLICATION NUMBER: US/09/787,986A
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Conus marmoreus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (12)
; OTHER INFORMATION: Xaa at position 12 is 4-hydroxy proline
US-09-787-986A-2

Query Match 83.6%; Score 61; DB 5; Length 13;
Best Local Similarity 83.3%; Pred. No. 0.0052;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVCCGXXLCHXC 13
||||| ||| |
Db 2 GVCCGYKLCHXC 13

RESULT 5

US-10-123-155-61
; Sequence 61, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 61
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-61

Query Match 61.6%; Score 45; DB 6; Length 2119;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GVCCGXXLCHXC 13
| ||| | |
Db 316 gacgtaacatc 327

RESULT 6

US-10-123-155-77
; Sequence 77, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 77
; LENGTH: 3567
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-77

Query Match 61.6%; Score 45; DB 6; Length 3567;

Best Local Similarity 50.0%; Pred. No. 79;

Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVC CGXXLCHXC 13

Db 2668 gaccgtatcttc 2679

RESULT 7

US-10-123-155-441

; Sequence 441, Application US/10123155

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330R1C30

; CURRENT APPLICATION NUMBER: US/10/123,155

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 441

; LENGTH: 654

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-123-155-441

Query Match 60.3%; Score 44; DB 6; Length 654;

Best Local Similarity 50.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVC CGXXLCHXC 13

Db 395 gtccggttcaac 406

RESULT 8

US-10-123-155-3

; Sequence 3, Application US/10123155

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 3
; LENGTH: 1395
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-3

Query Match 60.3%; Score 44; DB 6; Length 1395;

Best Local Similarity 50.0%; Pred. No. 49;

Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVC CGXXLCHXC 13

Db 158 gaccggttctac 169

RESULT 9

US-10-123-155-65

; Sequence 65, Application US/10123155

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330R1C30

; CURRENT APPLICATION NUMBER: US/10/123,155

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 65

; LENGTH: 1606

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-123-155-65

Query Match 60.3%; Score 44; DB 6; Length 1606;

Best Local Similarity 50.0%; Pred. No. 55;

Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVC CGXXLCHXC 13

Db 158 gaccggttctac 169

Db 422 gtccgatacgcc 433

```
RESULT 10
US-10-123-155-257
; Sequence 257, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 257
; LENGTH: 1606
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-257
```

Query Match 60.3%; Score 44; DB 6; Length 1606;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13
| | | | |
Db 422 gtccgatacgcc 433

```
RESULT 11
US-10-063-545-85
; Sequence 85, Application US/10063545
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,545
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 85
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-545-85
```

Query Match 60.3%; Score 44; DB 6; Length 1665;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13
| | | | |
Db 307 gaccgattccac 318

```
RESULT 12
US-10-063-685-85
; Sequence 85, Application US/10063685
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,685
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 85
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-685-85
```

Query Match 60.3%; Score 44; DB 6; Length 1665;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13
| | | | |
Db 307 gaccgattccac 318

```
RESULT 13
US-10-123-155-9
; Sequence 9, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
```

; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 9
 ; LENGTH: 2276
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-123-155-9

Query Match 60.3%; Score 44; DB 6; Length 2276;

Best Local Similarity 50.0%; Pred. No. 74;

Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13

Db 1747 gaccgaacctgc 1758

RESULT 14

US-10-123-155-79

; Sequence 79, Application US/10123155

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330RIC30

; CURRENT APPLICATION NUMBER: US/10/123,155

; CURRENT FILING DATE: 2002-04-15

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 79

; LENGTH: 2714

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-123-155-79

Query Match

60.3%; Score 44; DB 6; Length 2714;

Best Local Similarity 50.0%; Pred. No. 86;

Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13

Db 644 gcccgatacaac 655

RESULT 15

US-10-123-155-405

; Sequence 405, Application US/10123155

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330RIC30
 ; CURRENT APPLICATION NUMBER: US/10/123,155
 ; CURRENT FILING DATE: 2002-04-15
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 405
 ; LENGTH: 3819
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-123-155-405

Query Match 60.3%; Score 44; DB 6; Length 3819;

Best Local Similarity 50.0%; Pred. No. 11e+02;

Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13

Db 904 gaccggatcaac 915

Search completed: May 19, 2002, 08:04:48

Job time: 15749 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 08:00:18 ; Search time 324.64 Seconds
(without alignments)
14.095 Million cell updates/sec

Title: US-09-580-201A-2
Perfect score: 73
Sequence: 1 NGVCGXXLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	91.8	13	1	PCT-US00-01978-2
2	67	91.8	13	19	US-09-580-201A-2
3	66	90.4	61	1	PCT-US00-01978-9
4	66	90.4	61	19	US-09-580-201A-12
5	61	83.6	12	1	PCT-US00-01978-3
6	61	83.6	12	1	PCT-US00-01978-4
7	61	83.6	12	19	US-09-580-201A-3

8	61	83.6	12	19	US-09-580-201A-4	Sequence 4, Appli
9	49	67.1	60	1	PCT-US99-22853B-1309	Sequence 1309, Ap
10	45	61.6	24	11	US-08-775-765-8	Sequence 8, Appli
11	45	61.6	24	11	US-08-775-765-8	Sequence 8, Appli
12	45	61.6	1591	24	US-10-073-912-13	Sequence 13, Appli
13	44	60.3	41	26	US-60-162-248-1396	Sequence 1396, Ap
14	44	60.3	538	21	US-09-779-307-2	Sequence 2, Appli
15	44	60.3	538	21	US-09-779-307-11	Sequence 11, Appli
16	44	60.3	547	21	US-09-779-307-12	Sequence 12, Appli
17	44	60.3	547	21	US-09-779-307-13	Sequence 13, Appli
18	44	60.3	606	26	US-60-187-409-38	Sequence 38, Appli
19	43	58.9	92	1	PCT-US01-08656-8013	Sequence 8013, Ap
20	43	58.9	118	16	US-09-252-991A-24721	Sequence 24721, A
21	43	58.9	723	26	US-60-230-445-1241	Sequence 1241, Ap
22	43	58.9	766	1	PCT-US01-04098A-1921	Sequence 1921, Ap
23	43	58.9	1203	1	PCT-US01-08631-35830	Sequence 35830, A
24	43	58.9	1597	1	PCT-US01-08631-34348	Sequence 34348, A
25	43	58.9	1597	1	PCT-US01-08631-40612	Sequence 40612, A
26	43	58.9	1597	1	PCT-US01-08631-41171	Sequence 41171, A
27	43	58.9	1597	1	PCT-US01-08631-50261	Sequence 50261, A
28	43	58.9	1599	1	PCT-US01-08631-38243	Sequence 38243, A
29	43	58.9	1599	1	PCT-US01-08631-44809	Sequence 44809, A
30	43	58.9	1784	1	PCT-US01-08631-39507	Sequence 39507, A
31	43	58.9	1917	18	US-09-436-063-5	Sequence 5, Appli
32	43	58.9	1917	18	US-09-436-063C-5	Sequence 5, Appli
33	43	58.9	1917	20	US-09-627-650-5	Sequence 5, Appli
34	43	58.9	1917	20	US-09-627-650A-5	Sequence 5, Appli
35	43	58.9	1917	20	US-09-627-650B-5	Sequence 5, Appli
36	43	58.9	2572	1	PCT-US01-12836-26	Sequence 26, Appli
37	43	58.9	2572	19	US-09-559-001-26	Sequence 26, Appli
38	43	58.9	2572	21	US-09-728-403-26	Sequence 26, Appli
39	42	57.5	13	1	PCT-US00-01978-1	Sequence 1, Appli
40	42	57.5	59	26	US-60-162-357-938	Sequence 938, App
41	42	57.5	59	26	US-60-170-429-1072	Sequence 1072, Ap
42	42	57.5	59	26	US-60-178-306-1261	Sequence 1261, Ap
43	42	57.5	68	26	US-60-170-430-2088	Sequence 2088, Ap
44	42	57.5	68	26	US-60-177-646-2406	Sequence 2406, Ap
45	42	57.5	70	26	US-60-170-374-2636	Sequence 2636, Ap

ALIGNMENTS

RESULT 1
PCT-US00-01978-2
; Sequence 2, Application PC/TUS0001978
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Conotoxin Peptides
; FILE REFERENCE: ap-conotoxins
; CURRENT APPLICATION NUMBER: PCT/US00/01978
; CURRENT FILING DATE: 2000-01-28
; EARLIER APPLICATION NUMBER: US 60/173,298
; EARLIER FILING DATE: 1999-12-28
; EARLIER APPLICATION NUMBER: US 60/118,381
; EARLIER FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Conus marmoreus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (7)..(8)
; OTHER INFORMATION: Xaa at residue 7 is Tyr, mono-halo-Tyr,
; OTHER INFORMATION: di-halo-Tyr, O-sulpho-Tyr, O-Phospho-Tyr or
; OTHER INFORMATION: nitro-Tyr; Xaa at residue 8 is Lys, N-methyl-Lys,
; OTHER INFORMATION: N,N-dimethyl-Lys or N,N,N-trimethyl Lys
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (12)
; OTHER INFORMATION: Xaa at residue 12 is Pro or hydroxy-Pro.

PCT-US00-01978-2

Query Match 91.8%; Score 67; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
| | | | | | | | | | | | |
DB 1 NGVCCGXXLCHXC 13

RESULT 2

US-09-580-201A-2
Sequence 2, Application US/09580201A

GENERAL INFORMATION:
APPLICANT: McIntosh, J. Michael
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Corpuz, Gloria P.
APPLICANT: Jones, Robert M.
APPLICANT: Garrett, James E.

TITLE OF INVENTION: Conotoxin Peptides
FILE REFERENCE: Conotoxin Peptides
CURRENT APPLICATION NUMBER: US/09/580, 201A
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 60/173,298
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: US 60/118,381
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: US 09/493,143
PRIOR FILING DATE: 2000-01-28

NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 13

TYPE: PRT

ORGANISM: Conus marmoreus

FEATURE:

NAME/KEY: PEPTIDE
LOCATION: (7)..(8)

OTHER INFORMATION: Xaa at residue 7 is Tyr, mono-halo-Tyr.
OTHER INFORMATION: di-halo-Tyr, O-sulpho-Tyr, O-Phospho-Tyr or
OTHER INFORMATION: nitro-Tyr; Xaa at residue 8 is Lys, N-methyl-Lys,
OTHER INFORMATION: N,N-dimethyl-Lys or N,N,N-trimethyl Lys

NAME/KEY: PEPTIDE

LOCATION: (12)

OTHER INFORMATION: Xaa at residue 12 is Pro or hydroxy-Pro.

US-09-580-201A-2

Query Match 91.8%; Score 67; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
| | | | | | | | | | | | |
DB 1 NGVCCGXXLCHXC 13

RESULT 3

PCT-US00-01978-9

Sequence 9, Application PC/TUS00001978
GENERAL INFORMATION:

APPLICANT: University of Utah Research Foundation
TITLE OF INVENTION: Conotoxin Peptides
FILE REFERENCE: ap-conotoxins

CURRENT APPLICATION NUMBER: PCT/US00/01978

CURRENT FILING DATE: 2000-01-28

EARLIER APPLICATION NUMBER: US 60/173,298

EARLIER FILING DATE: 1999-12-28

EARLIER APPLICATION NUMBER: US 60/118,381

EARLIER FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 61
TYPE: PRT
ORGANISM: Conus marmoreus
PCT-US00-01978-9

Query Match 90.4%; Score 66; DB 1; Length 61;
Best Local Similarity 76.9%; Pred. No. 0.71;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
| | | | | | | | | | | | |
DB 49 NGVCCGXXLCHXC 61

RESULT 4

US-09-580-201A-12
Sequence 12, Application US/09580201A

GENERAL INFORMATION:
APPLICANT: McIntosh, J. Michael
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Corpuz, Gloria P.
APPLICANT: Jones, Robert M.
APPLICANT: Garrett, James E.

TITLE OF INVENTION: Conotoxin Peptides

FILE REFERENCE: Conotoxin Peptides

CURRENT APPLICATION NUMBER: US/09/580, 201A

CURRENT FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 60/173,298

PRIOR FILING DATE: 1999-12-28

PRIOR APPLICATION NUMBER: US 60/118,381

PRIOR FILING DATE: 1999-01-29

PRIOR APPLICATION NUMBER: US 09/493,143

PRIOR FILING DATE: 2000-01-28

NUMBER OF SEQ ID NOS: 20

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 12

LENGTH: 61

TYPE: PRT

ORGANISM: Conus marmoreus

US-09-580-201A-12

Query Match 90.4%; Score 66; DB 19; Length 61;
Best Local Similarity 76.9%; Pred. No. 0.71;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
| | | | | | | | | | | | |
DB 49 NGVCCGXXLCHXC 61

RESULT 5

PCT-US00-01978-3

Sequence 3, Application PC/TUS00001978

GENERAL INFORMATION:

APPLICANT: University of Utah Research Foundation

TITLE OF INVENTION: Conotoxin Peptides

FILE REFERENCE: ap-conotoxins

CURRENT APPLICATION NUMBER: PCT/US00/01978

CURRENT FILING DATE: 2000-01-28

EARLIER APPLICATION NUMBER: US 60/173,298

EARLIER FILING DATE: 1999-12-28

EARLIER APPLICATION NUMBER: US 60/118,381

EARLIER FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 9

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 12

```

; TYPE: PRT
; ORGANISM: Conus marmoreus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (6)..(7)
; OTHER INFORMATION: xaa at residue 6 is Tyr, mono-halo-Tyr,
; OTHER INFORMATION: di-halo-Tyr, O-sulpho-Tyr, O-phospho-Tyr or
; OTHER INFORMATION: nitro-Tyr; xaa at residue 7 is Lys, N-methyl-Lys,
; OTHER INFORMATION: N,N-dimethyl-Lys or N,N,N-trimethyl Lys
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (11)
; OTHER INFORMATION: xaa at residue 11 is Pro or hydroxy-Pro
PCT-US00-01978-3

Query Match      83.6%; Score 61; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2  GVCGXXLCHXC 13
    | | | | | | | |
Db   1  GVCGXXLCHXC 12

RESULT 6
PCT-US00-01978-4
; Sequence 4, Application PC/TUS0001978
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; FILE REFERENCE: Conotoxin Peptides
; TITLE OF INVENTION: Conotoxin Peptides
; CURRENT APPLICATION NUMBER: PCT/US00/01978
; EARLIER FILING DATE: 2000-01-28
; EARLIER APPLICATION NUMBER: US 60/173,298
; EARLIER FILING DATE: 1999-12-28
; EARLIER APPLICATION NUMBER: US 60/118,381
; EARLIER FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: unknown Conus
; OTHER INFORMATION: species
; NAME/KEY: PEPTIDE
; LOCATION: (6)..(7)
; OTHER INFORMATION: xaa at residue 6 is Tyr, mono-halo-Tyr,
; OTHER INFORMATION: di-halo-Tyr, O-sulpho-Tyr, O-phospho-Tyr or
; OTHER INFORMATION: nitro-Tyr; xaa at residue 7 is Lys, N-methyl-Lys,
; OTHER INFORMATION: N,N-dimethyl-Lys or N,N,N-trimethyl Lys.
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (11)
; OTHER INFORMATION: xaa at residue 11 is Pro or hydroxy-Pro
PCT-US00-01978-4

Query Match      83.6%; Score 61; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2  GVCGXXLCHXC 13
    | | | | | | | |
Db   1  GVCGXXLCHXC 12

```

```

; GENERAL INFORMATION:
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Corpuz, Gloria P.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; TITLE OF INVENTION: Conotoxin Peptides
; FILE REFERENCE: Conotoxin Peptides
; CURRENT APPLICATION NUMBER: US/09/580,201A
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/173,298
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/118,381
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: US 09/493,143
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Conus marmoreus
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (6)..(7)
; OTHER INFORMATION: xaa at residue 6 is Tyr, mono-halo-Tyr,
; OTHER INFORMATION: di-halo-Tyr, O-sulpho-Tyr, O-phospho-Tyr or
; OTHER INFORMATION: nitro-Tyr; xaa at residue 7 is Lys, N-methyl-Lys,
; OTHER INFORMATION: N,N-dimethyl-Lys or N,N,N-trimethyl Lys
; NAME/KEY: PEPTIDE
; LOCATION: (11)
; OTHER INFORMATION: xaa at residue 11 is Pro or hydroxy-Pro
US-09-580-201a-3

Query Match      83.6%; Score 61; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2  GVCGXXLCHXC 13
    | | | | | | | |
Db   1  GVCGXXLCHXC 12

RESULT 8
US-09-580-201A-4
; Sequence 4, Application US/09580201A
; GENERAL INFORMATION:
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Corpuz, Gloria P.
; APPLICANT: Jones, Robert M.
; APPLICANT: Garrett, James E.
; TITLE OF INVENTION: Conotoxin Peptides
; FILE REFERENCE: Conotoxin Peptides
; CURRENT APPLICATION NUMBER: US/09/580,201A
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/173,298
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/118,381
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: US 09/493,143
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: unknown Conus

```

```

RESULT 7
US-09-580-201A-3
; Sequence 3, Application US/09580201A

```

```
; OTHER INFORMATION: species
; NAME/KEY: PEPTIDE
; LOCATION: (6)..(7)
; OTHER INFORMATION: Xaa at residue 6 is Tyr, mono-halo-Tyr,
; OTHER INFORMATION: di-halo-Tyr, O-sulpho-Tyr, O-Phospho-Tyr or
; OTHER INFORMATION: nitro-Tyr; Xaa at residue 8 is Lys, N-methyl-Lys,
; OTHER INFORMATION: N,N-dimethyl-Lys or N,N,N-trimethyl Lys.
; NAME/KEY: PEPTIDE
; LOCATION: (11)
; OTHER INFORMATION: Xaa at residue 11 is Pro or hydroxy-Pro
US-09-580-201a-4

Query Match      83.6%; Score 61; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVC CGXXLCHXC 13
Db 1 GVC CGXXLCHXC 12

RESULT 9
PCT-US99-22853B-1309
; Sequence 1309, Application PC/TUS9922853B
; GENERAL INFORMATION:
; APPLICANT: Ceres, Inc.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-0569F(PC)
; CURRENT APPLICATION NUMBER: PCT/US99/22853B
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 3938
; SOFTWARE: MS Word 97
; SEQ ID NO 1309
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: LOCATION 1..60, Ceres Seq. ID 1688425
; NAME/KEY: UNSURE
; LOCATION: (1)..(60)
; OTHER INFORMATION: any Xaa = any amino acid, unknown or other
PCT-US99-22853B-1309

Query Match      67.1%; Score 49; DB 1; Length 60;
Best Local Similarity 53.8%; Pred. No. 85;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NGVCCGXXLCHXC 13
Db 10 HGVCCHWSCHFC 22

RESULT 10
US-08-775-765-8
; Sequence 8, Application US/08775765A
; GENERAL INFORMATION:
; APPLICANT: KELLEHER, Dermot
; APPLICANT: WINDLE, Henry
; APPLICANT: BYRNE, William
; APPLICANT: McMANUS, Ross
; TITLE OF INVENTION: HELICOBACTER PROTEINS AND VACCINES
; FILE REFERENCE: P60777.SEQ
; CURRENT APPLICATION NUMBER: US/08/775,765A
; CURRENT FILING DATE: 1996-12-31
; EARLIER APPLICATION NUMBER: IR 94 0538
; EARLIER FILING DATE: 1994-07-01
; EARLIER APPLICATION NUMBER: IR 95 0249
; EARLIER FILING DATE: 1995-04-06
; EARLIER APPLICATION NUMBER: PCT/IE 95/00036
; EARLIER FILING DATE: 1995-07-03

; OTHER INFORMATION: any Xaa = any amino acid, unknown or other
PCT-US99-22853B-1309

Query Match      61.6%; Score 45; DB 11; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GVC CGXXLCHXC 13
Db 6 GGC CGTYTCVTC 17

RESULT 11
US-08-775-765-8
; Sequence 8, Application US/08775765C
; GENERAL INFORMATION:
; APPLICANT: Kelleher, Dermot
; APPLICANT: Windle, Henry
; APPLICANT: Byrne, William
; APPLICANT: McManus, Ross
; TITLE OF INVENTION: Helicobacter Proteins and Vaccines
; FILE REFERENCE: 08/775 765
; CURRENT APPLICATION NUMBER: US/08/775,765C
; CURRENT FILING DATE: 1996-12-31
; EARLIER APPLICATION NUMBER: IE 94 0538
; EARLIER FILING DATE: 1994-07-01
; EARLIER APPLICATION NUMBER: IE 95 0249
; EARLIER FILING DATE: 1995-04-06
; EARLIER APPLICATION NUMBER: PCT/IE 95/00036
; EARLIER FILING DATE: 1995-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-08-775-765-8

Query Match      61.6%; Score 45; DB 11; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GVC CGXXLCHXC 13
Db 6 GGC CGTYTCVTC 17

RESULT 12
US-10-073-912-13
; Sequence 13, Application US/10073912
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT234CI
; CURRENT APPLICATION NUMBER: US/10/073,912
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1591
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (33)
```

; OTHER INFORMATION: n equals a,t,g, or c
US-10-073-912-13

Query Match 61.6%; Score 45; DB 24; Length 1591;
Best Local Similarity 50.0%; Pred. No. 2.2e+03;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13
| | | | |
Db 1010 gaccgaaccatc 1021

RESULT 13
US-60-162-248-1396
; Sequence 1396, Application US/60162248
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000123
; CURRENT APPLICATION NUMBER: US/60/162,248
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 1476
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1396
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Human
US-60-162-248-1396

Query Match 60.3%; Score 44; DB 26; Length 41;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC---GXLLCHXC 13
| | | | |
Db 11 GCCCFPLDGHLLCHGC 26

RESULT 14
US-09-779-307-2
; Sequence 2, Application US/09779307
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-307-2

Query Match 60.3%; Score 44; DB 21; Length 538;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;

Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC---GXLLCHXC 13
| | | | |
Db 508 GCCCFPLDGHLLCHGC 523

RESULT 15
US-09-779-307-11
; Sequence 11, Application US/09779307
; GENERAL INFORMATION:
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Prayaga, Sudhirdas
; TITLE OF INVENTION: Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-662 US
; CURRENT APPLICATION NUMBER: US/09/779,307
; CURRENT FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/180,880
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,044
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/181,656
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/182,795
; PRIOR FILING DATE: 2000-02-15
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-779-307-11

Query Match 60.3%; Score 44; DB 21; Length 538;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC---GXLLCHXC 13
| | | | |
Db 508 GCCCFPLDGHLLCHGC 523

Search completed: May 19, 2002, 08:00:18
Job time: 21700 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 04:44:12 ; Search time 42.19 seconds
(without alignments)
7.526 Million cell updates/sec

Title: US-09-580-201A-2
Perfect score: 73
Sequence: 1 NGVCGXXLCHXC 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
* score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match		DB ID	Description
		Match	Length		
1	41.5	56.8	768	1	US-08-454-455-4
2	41	56.2	574	1	US-08-049-473-2
3	41	56.2	574	1	US-08-312-648-2
4	41	56.2	574	5	PCT-US94-04190-2
5	38.5	52.7	26	4	US-09-073-407-9
6	38.5	52.7	26	4	US-09-073-407-12
7	38	52.1	485	2	US-08-724-394A-8
8	38	52.1	801	1	US-07-906-349A-6
9	38	52.1	1345	2	US-08-977-767-3
10	38	52.1	1417	4	US-08-900-230-3
11	37.5	51.4	26	4	US-09-073-407-11
12	36.5	50.0	25	1	US-08-084-848A-4
13	36.5	50.0	25	1	US-08-458-499-4
14	36.5	50.0	26	4	US-09-073-407-10
15	36.5	50.0	1810	5	PCT-US95-11684-4
16	36	49.3	18	1	US-08-084-848A-6
17	36	49.3	18	1	US-08-458-499-6
18	36	49.3	38	1	US-07-609-716-53
19	36	49.3	44	4	US-09-046-894-41
20	36	49.3	44	4	US-09-046-894-42
21	36	49.3	441	2	US-08-491-835-4
22	36	49.3	441	3	US-08-946-092A-4
23	36	49.3	441	4	US-09-172-062-4
24	36	49.3	441	5	PCT-US94-00685-4
25	35.5	48.6	22	1	US-08-423-399B-27
26	35.5	48.6	25	1	US-08-084-848A-1
27	35.5	48.6	25	1	US-08-084-848A-5

28	35.5	48.6	25	1	US-08-458-499-1	Sequence 1, Appl
29	35.5	48.6	25	1	US-08-458-499-5	Sequence 5, Appl
30	35.5	48.6	30	1	US-08-137-800-33	Sequence 33, Appl
31	35.5	48.6	30	1	US-08-477-383-33	Sequence 33, Appl
32	35.5	48.6	30	1	US-08-480-174-33	Sequence 33, Appl
33	35.5	48.6	30	1	US-08-480-750-33	Sequence 33, Appl
34	35.5	48.6	769	1	US-08-454-455-6	Sequence 6, Appl
35	35.5	48.6	1070	4	US-08-697-954-2	Sequence 2, Appl
36	35	47.9	41	5	PCT-US96-01720-7	Sequence 7, Appl
37	35	47.9	47	2	US-08-691-814B-20	Sequence 20, Appl
38	35	47.9	111	2	US-08-485-937-4	Sequence 4, Appl
39	35	47.9	111	2	US-08-373-215-4	Sequence 4, Appl
40	35	47.9	111	5	PCT-US93-06552-4	Sequence 4, Appl
41	35	47.9	141	2	US-08-485-937-1	Sequence 1, Appl
42	35	47.9	141	2	US-08-373-215-1	Sequence 1, Appl
43	35	47.9	141	5	PCT-US93-06552-1	Sequence 1, Appl
44	35	47.9	199	1	US-08-264-534-4	Sequence 4, Appl
45	35	47.9	199	1	US-08-454-455-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-454-455-4
; Sequence 4, Application US/08454455
; Patent No. 5635601
; GENERAL INFORMATION:
; APPLICANT: Moyle, Matthew
; APPLICANT: McLean, John W.
; TITLE OF INVENTION: NOVEL BETA INTEGRIN SUBUNIT
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,455
; FILING DATE: 30-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/193989
; FILING DATE: 09-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/004142
; FILING DATE: 13-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/670607
; FILING DATE: 14-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: P0699C2D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 768 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-454-455-4

```
Query Match      56.8%; Score 41.5; DB 1; Length 768;
Best Local Similarity 72.7%; Pred. No. 1.2e+02;
Matches      8; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy      2 GVC-CGXXLCH 11
Db      524 GVCVCGKCLCH 534

RESULT      2
US-08-049-473-2
; Sequence 2, Application US/08049473
; Patent No. 5386021
; GENERAL INFORMATION:
; APPLICANT: Moss, Joel
; APPLICANT: Mishima, Koichi
; APPLICANT: Nightingale, Maria
; APPLICANT: Tsuchiya, Mikako
; TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,473
; FILING DATE: 19930419
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael L.
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH050.001CP1
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 574 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-049-473-2

Query Match      56.28; Score 41; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches      8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

Qy      2 GVC-----CGXXLCHXC 13
Db      32 GVCEDVFSLQDKVPRLLLCGHTVCHDC 59

RESULT      4
PCT-US94-04190-2
; Sequence 2, Application PC/TUS9404190
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary, Department
; APPLICANT: of Health and Human Services
; TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
; TITLE OF INVENTION: PROTEIN WITH AN ADP-RYBOSYLATION FACTOR DOMAIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04190
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael L.
```

```
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,648
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/049,473
; FILING DATE: 19-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael L.
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH050.001DVI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 574 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-312-648-2

Query Match      56.2%; Score 41; DB 1; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches      8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

Qy      2 GVC-----CGXXLCHXC 13
Db      32 GVCEDVFSLQDKVPRLLLCGHTVCHDC 59

RESULT      4
PCT-US94-04190-2
; Sequence 2, Application PC/TUS9404190
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary, Department
; APPLICANT: of Health and Human Services
; TITLE OF INVENTION: A MAMMALIAN GUANIN NUCLEOTIDE BINDING
; TITLE OF INVENTION: PROTEIN WITH AN ADP-RYBOSYLATION FACTOR DOMAIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04190
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael L.
```


REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH050.001QPC
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-04190-2

Query Match 56.2%; Score 41; DB 5; Length 574;
Best Local Similarity 28.6%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 16; Gaps 1;

QY 2 GVC-----CGXXLCHXC 13
||| |||
Db 32 GVCEDVFSLQGVPRLLLCGHTVCHDC 59

RESULT 5
US-09-073-407-9
Sequence 9, Application US/09073407
Patent No. 6232065
GENERAL INFORMATION:
APPLICANT: Robinson, Daniel R.
APPLICANT: Kung, Hsing-Jien
TITLE OF INVENTION: ANALYSIS OF GENE FAMILY EXPRESSION
FILE REFERENCE: CASE-03147
CURRENT APPLICATION NUMBER: US/09/073,407
CURRENT FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-073-407-9

Query Match 52.7%; Score 38.5; DB 4; Length 26;
Best Local Similarity 46.2%; Pred. No. 1.7;
Matches 6; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 NGVCCGXXLCHXC 13
||| |||
Db 15 NGMCCAAA-CRTC 26

RESULT 6
US-09-073-407-12
Sequence 12, Application US/09073407
Patent No. 6232065
GENERAL INFORMATION:
APPLICANT: Robinson, Daniel R.
APPLICANT: Kung, Hsing-Jien
TITLE OF INVENTION: ANALYSIS OF GENE FAMILY EXPRESSION
FILE REFERENCE: CASE-03147
CURRENT APPLICATION NUMBER: US/09/073,407
CURRENT FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-073-407-12

Query Match 52.7%; Score 38.5; DB 4; Length 26;
Best Local Similarity 46.2%; Pred. No. 1.7;
Matches 6; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 NGVCCGXXLCHXC 13
||| |||
Db 15 NGMCCATA-CRTC 26

RESULT 7
US-08-724-394A-8
Sequence 8, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 1..485
OTHER INFORMATION: /note= "RoRet"
US-08-724-394A-8

Query Match 52.1%; Score 38; DB 2; Length 485;
Best Local Similarity 55.6%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CGXXLCHXC 13
||| |||
Db 31 CGHSYCHLC 39

```
RESULT 8
US-07-906-349A-6
; Sequence 6, Application US/07906349A
; Patent No. 5434064
; GENERAL INFORMATION:
; APPLICANT: Schlensing, Joseph
; APPLICANT: Skolnik, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AN
; TITLE OF INVENTION: TARGET PROTEINS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,349A
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/643,237
; FILING DATE: 18-JAN-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 801 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-906-349A-6

Query Match 52.1%; Score 38; DB 1; Length 801;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCGXXLCHXC 13
| | | | |
Db 339 GGCGGTCACC 350

RESULT 9
US-07-977-767-3
; Sequence 3, Application US/08977767
; Patent No. 5972684
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Greenwald, Sara
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: CARBONIC ANHYDRASE VIII
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

Query Match 52.1%; Score 38; DB 1; Length 801;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCGXXLCHXC 13
| | | | |
Db 339 GGCGGTCACC 350
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,767
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0423 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1345 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1532042
US-08-977-767-3

Query Match 52.1%; Score 38; DB 2; Length 1345;
Best Local Similarity 50.0%; Pred. No. 5.6e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXLCHXC 13
| | | | |
Db 1089 CCGTAACGC 1098

RESULT 10
US-08-900-230-3
; Sequence 3, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/900,230
; APPLICATION NUMBER: 23-JUL-1997
; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
```

; LENGTH: 1417 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: NO
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-900-230-3

Query Match 52.1%; Score 38; DB 4; Length 1417;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GVCGGXXLCHXC 13
| | | | |
Db 958 GCGCGTTCGCC 969

RESULT 11

US-09-073-407-11
; Sequence 11, Application US/09073407
; Patent No. 6232065
; GENERAL INFORMATION:
; APPLICANT: Robinson, Daniel R.
; APPLICANT: Kung, Hsing-Jien
; TITLE OF INVENTION: ANALYSIS OF GENE FAMILY EXPRESSION
; FILE REFERENCE: CASE-03147
; CURRENT APPLICATION NUMBER: US/09/073,407
; CURRENT FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-073-407-11

Query Match 51.4%; Score 37.5; DB 4; Length 26;
Best Local Similarity 46.2%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 NGVCCGXXLCHXC 13
| | | | |
Db 15 NGMCCAGA-CRTC 26

RESULT 12

US-08-084-848A-4
; Sequence 4, Application US/08084848A
; Patent No. 5432155
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Rivier, Jean E. F.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Abogadie, Fe
; APPLICANT: Hopkins, Chris E.
; APPLICANT: Dykert, John
; APPLICANT: Torres, Josep L.
; TITLE OF INVENTION: Conotoxins I
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 South LaSalle Street, Suite 900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/084,848A
; FILING DATE: June 29, 1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Schumann, James J.
; REGISTRATION NUMBER: 20856
; REFERENCE/DOCKET NUMBER: 52511
; TELEPHONE: (619)552-1311
; TELEFAX: (619)552-0095
; TELEX: 20 6566 PATLAW CGO
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-084-848A-4

Query Match 50.0%; Score 36.5; DB 1; Length 25;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 4 CCGX---XLCHXC 13
| | | | |
Db 2 CCGSYXNACHXC 14

RESULT 13

US-08-458-499-4
; Sequence 4, Application US/08458499
; Patent No. 5700778
; GENERAL INFORMATION:
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Rivier, Jean E. F.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Abogadie, Fe
; APPLICANT: Hopkins, Chris E.
; APPLICANT: Dykert, John
; APPLICANT: Torres, Josep L.
; TITLE OF INVENTION: Conotoxins I
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 South LaSalle Street, Suite 900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,499
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/084,848
; FILING DATE: June 29, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Schumann, James J.
; REGISTRATION NUMBER: 20856
; REFERENCE/DOCKET NUMBER: 52511
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)552-1311
; TELEFAX: (619)552-0095

; TELEX: 20 6566 PATLAW CGO
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-458-499-4

Query Match 50.0%; Score 36.5; DB 1; Length 25;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

Qy 4 CCGX---XLCHXC 13
||| ||||
Db 2 CCGSYXNAACHXC 14

RESULT 14
US-09-073-407-10
; Sequence 10, Application US/09073407
; Patent No. 6232065
; GENERAL INFORMATION:
; APPLICANT: Robinson, Daniel R.
; APPLICANT: Kung, Hsing-Jien
; TITLE OF INVENTION: ANALYSIS OF GENE FAMILY EXPRESSION
; FILE REFERENCE: CASE-03147
; CURRENT APPLICATION NUMBER: US/09/073,407
; CURRENT FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-073-407-10

Query Match 50.0%; Score 36.5; DB 4; Length 26;
Best Local Similarity 46.2%; Pred. No. 31;
Matches 6; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 1 NGVCCGXXLCHXC 13
||| |||
Db 15 NGMCCACA-CRTC 26

RESULT 15
PCT-US95-11684-4
; Sequence 4, Application PC/TUS9511684
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE
; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING SAME
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/11684
; FILING DATE: 14-SEP-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,359
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Logan, April C.
; REGISTRATION NUMBER: 33,950
; REFERENCE/DOCKET NUMBER: BEC0019P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1810 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-11684-4

Query Match 50.0%; Score 36.5; DB 5; Length 1810;
Best Local Similarity 37.5%; Pred. No. 1.1e+03;
Matches 9; Conservative 1; Mismatches 3; Indels 11; Gaps 2;

Qy 1 NGVC-----CGXXLC-HXC 13
||| ||| |||
Db 235 DGVCVCFEGYGTGDCGELCPHGC 258

Search completed: May 19, 2002, 04:44:13
Job time: 12551 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 04:43:16 ; Search time 86.36 Seconds
(without alignments)
16.720 Million cell updates/sec

Title: US-09-580-201A-2

Perfect score: 73

Sequence: 1 NGVCGXXLCHXC 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SID55/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	90.4	13	21 AAB08013	Generic formula fo
2	66	90.4	13	21 AAB08017	Amino acid sequenc
3	66	90.4	13	21 AAY92229	Chi-conotoxin pept
4	66	90.4	61	21 AAB08016	Amino acid sequenc
5	66	90.4	61	21 AAY92231	Chi-conotoxin, chi
6	60	82.2	12	21 AAB08014	Generic formula fo
7	60	82.2	12	21 AAB08015	Generic formula fo
8	60	82.2	12	21 AAB08018	Amino acid sequenc
9	60	82.2	12	21 AAB08019	Amino acid sequenc
10	60	82.2	13	21 AAY92230	Chi-conotoxin pept
11	44	60.3	538	22 AAU06111	Novel human polype

12	43	58.9	93	22 AAU31406	Novel human secret
13	43	58.9	766	22 AAM79259	Human protein SEQ
14	43	58.9	1203	22 ABG05471	Novel human diagno
15	43	58.9	1597	22 ABG03989	Novel human diagno
16	43	58.9	1597	22 ABG10253	Novel human diagno
17	43	58.9	1597	22 ABG10812	Novel human diagno
18	43	58.9	1597	22 ABG19902	Novel human diagno
19	43	58.9	1599	22 ABG07884	Novel human diagno
20	43	58.9	1599	22 ABG14450	Novel human diagno
21	43	58.9	1784	22 ABG09148	Novel human diagno
22	42	57.5	369	22 ABG15487	Novel human diagno
23	42	57.5	865	22 ABG25368	Novel human diagno
24	42	57.5	1129	22 AAB62174	Human SV protein.
25	42	57.5	1169	22 AAB62173	Human Rabkayrin-5
26	42	57.5	1170	22 AAM40022	Human polypeptide
27	41.5	56.8	768	13 AAR27683	Rabbit beta-8 inte
28	41	56.2	13	21 AAB08012	Generic formula fo
29	41	56.2	60	20 AAY27648	Human secreted pro
30	41	56.2	164	21 AAG01721	Human secreted pro
31	41	56.2	360	22 AAU64532	Propionibacterium
32	41	56.2	437	20 AAY32502	Human parkin gene
33	41	56.2	465	20 AAY32501	Human parkin gene
34	41	56.2	574	15 AAR66034	Rat ARD 1. Rattus
35	41	56.2	574	15 AAR66033	Human ARD 1. Homo
36	41	56.2	919	22 AAG84970	Shrimp white spot
37	40	54.8	73	19 AAY20881	Human presenilin I
38	40	54.8	344	22 AAB67521	Amino acid sequenc
39	40	54.8	451	22 AAB67533	Amino acid sequenc
40	40	54.8	464	22 AAB67517	Amino acid sequenc
41	40	54.8	464	22 AAB67531	Amino acid sequenc
42	40	54.8	464	22 AAB67532	Amino acid sequenc
43	40	54.8	1023	22 AAB11436	D. discoideum ster
44	40	54.8	1231	22 ABB59738	Drosophila melanog
45	39	53.4	55	22 ABB66559	Drosophila melanog

ALIGNMENTS

RESULT	1
AAB08013	
ID	AAB08013 standard; peptide; 13 AA.
XX	AC
XX	AAB08013:
DT	14-NOV-2000 (first entry)
XX	Generic formula for conotoxin peptide Marl.
DE	
XX	Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
KW	
XX	Conus marmoreus.
OS	
XX	
PH	Key Location/Qualifiers
FT	Misc-difference 7 /label= Tyr, Xaa
FT	/note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
FT	O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT	
FT	Misc-difference 8 /label= Lys, Xaa
FT	/note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT	N,N,N-trimethyl Lys"
FT	
FT	Misc-difference 12 /note= "Optionally hydroxy-Pro"
FT	
XX	
XX	WC200004769-A1.
PN	
XX	
XX	03-AUG-2000.
PD	
XX	
XX	28-JAN-2000; 2000WO-US01978.
PF	
XX	
XX	29-JAN-1999; 99US-0118381.
PR	
PR	28-DEC-1999; 99US-0173343.

XX PA (UTAH) UNIV UTAH RES FOUND.
 XX PI McIntosh JM, Olivera BM, Cruz LJ;
 XX XX
 DR WPI; 2000-476222/41.
 XX PT Purified ap-conotoxin derived from cone snail venom for use as an
 XX PT analgesic -
 XX XX
 PS Claim 2; Page 19; 29pp; English.
 XX CC The present sequence represents an ap-conotoxin peptide, designated
 CC Marl. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 XX XX
 SQ Sequence 13 AA;
 Query Match 90.4%; Score 66; DB 21; Length 13;
 Best Local Similarity 92.3%; Pred. No. 0.014;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 NGVCCGXXLCHXC 13
 | | | | | | | | | | | | |
 Db 1 ngvccgxxlchpc 13
 RESULT 2
 AAB08017
 ID AAB08017 standard; peptide; 13 AA.
 XX AC AAB08017;
 XX DT 14-NOV-2000 (first entry)
 XX DE Amino acid sequence of the conotoxin peptide Marl.
 XX KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
 XX OS Conus marmoreus.
 XX FH Key Location/Qualifiers
 FT Modified-site 12
 FT FT /note= "hydroxy-Pro"
 XX PN W0200044769-A1.
 XX PD 03-AUG-2000.
 XX PF 28-JAN-2000; 2000W0-US01978.
 XX PR 29-JAN-1999; 99US-0118381.
 XX PR 28-DEC-1999; 99US-0173343.
 XX PA (UTAH) UNIV UTAH RES FOUND.
 XX PI McIntosh JM, Olivera BM, Cruz LJ;
 XX XX
 DR WPI; 2000-476222/41.
 XX PT Purified ap-conotoxin derived from cone snail venom for use as an
 XX PT analgesic -
 XX XX
 PS Claim 13; Page -; 29pp; English.
 XX CC The present sequence represents an ap-conotoxin peptide, designated
 CC Marl. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.

XX SQ Sequence 13 AA;
 Query Match 90.4%; Score 66; DB 21; Length 13;
 Best Local Similarity 76.9%; Pred. No. 0.014;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 NGVCCGXXLCHXC 13
 | | | | | | | | | | | | |
 Db 1 ngvccgxxlchpc 13
 RESULT 3
 AAY92229
 ID AAY92229 standard; peptide; 13 AA.
 XX AC AAY92229;
 XX DT 10-AUG-2000 (first entry)
 XX DE Chi-conotoxin peptide, chi-MrIA.
 XX KW chi-conotoxin; chi-MrIA; cone snail; inhibitor; amine transporter;
 KW neuronal; noradrenaline transporter; urinary tract disorder; analgesic;
 KW antiarrhythmic; cardiac; antidepressant; anxiolytic; anti-inflammatory.
 XX OS Conus marmoreus.
 XX FH Key Location/Qualifiers
 FT Misc-difference 12
 FT FT /label= 4Hyp
 FT FT /note= "4-hydroxyproline"
 XX PN W0200020444-A1.
 XX PD 13-APR-2000.
 XX PF 01-OCT-1999; 99W0-AU00844.
 XX PR 02-OCT-1998; 98AU-0006274.
 XX PA (UYQU) UNIV QUEENSLAND.
 XX PI Lewis RJ, Alewood PF, Sharpe IA;
 XX PR WPI; 2000-303738/26.
 XX PT Isolated, synthetic or recombinant chi-conotoxin peptide capable of
 PT inhibiting neuronal amine transporter used for treatment or prophylaxis
 PT of urinary or cardiovascular conditions, mood disorders, or
 PT treatment/control of pain/inflammation
 XX PS Claim 3; Page 33; 47pp; English.
 XX CC This conotoxin, chi-MrIA, is a member of a new class of conotoxins,
 CC designated chi-conotoxin. It was isolated from the venom of the mollusc
 CC hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the
 CC neuronal amine transporters, especially the neuronal noradrenaline
 CC transporter. Inhibitors of noradrenaline re-uptake which have a
 CC negligible anti-cholinergic effect are particularly useful in the
 CC treatment of lower urinary tract disorders. Chi-MrIA (0.1 nM-1 micro M)
 CC inhibited the accumulation of radiolabeled noradrenaline in a
 CC concentration-dependent manner, with a log IC-50 value of -8.17 plus or
 CC minus 0.0275 (n = 4). The concentration of chi-MrIA required to inhibit
 CC the accumulation by 50 percent was found to be approximately 7 nM. This
 CC concentration is approximately one order of magnitude lower than that
 CC needed for desipramine to produce the same effect. The peptides are
 CC useful for the treatment or prophylaxis of urinary or cardiovascular
 CC conditions or diseases (arrhythmia or coronary heart failure) or mood
 CC disorders (depression, anxiety or cravings), or the treatment or control
 CC of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 CC pain).

XX Sequence 13 AA;

SQ

Query Match 90.4%; Score 66; DB 21; Length 13;
Best Local Similarity 76.9%; Pred. No. 0.014;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
DB 1 ngvcgkklchpc 13

RESULT 4
ID AAB08016
XZ AAB08016 standard; Protein; 61 AA.
XX
AC AAB08016;
XZ 14-NOV-2000 (first entry)
XZ
XZ Amino acid sequence of the conotoxin Marl propeptide.
XZ
KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Marl.
XZ
OS Conus marmoreus.
XZ
PN WO200044769-A1.
XZ
PD 03-AUG-2000.
XZ
XZ 28-JAN-2000; 2000WO-US01978.
XZ
XZ 29-JAN-1999; 99US-0118381.
XZ 28-DEC-1999; 99US-0173343.
XZ (UTAH) UNIV UTAH RES FOUND.
XZ
XZ McIntosh JW, Olivera BW, Cruz LJ;
XZ WPI; 2000-476222/41.
XZ N-PSDB; AAA63513.
XZ
XZ Purified ap-conotoxin derived from cone snail venom for use as an
XZ analgesic -
XZ
XZ Claim 23; Page 13-14; 29pp; English.
XZ
XZ The present sequence represents a Marl propeptide. Marl is an
XZ ap-conotoxin peptide. Conotoxins are naturally available in minute
XZ amounts in the venom of cone snails. The peptides have analgesic
XZ activity. The peptides are used to treat or prevent pain.
XZ
SQ Sequence 61 AA;

Query Match 90.4%; Score 66; DB 21; Length 61;
Best Local Similarity 76.9%; Pred. No. 0.048;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
DB 49 ngvcgkklchpc 61

RESULT 5
ID AAY92231
XZ AAY92231 standard; Protein; 61 AA.
XZ
AC AAY92231;
XZ
XZ 10-AUG-2000 (first entry)
XZ
XZ

DE Chi-conotoxin, chi-MrIA, leader and mature peptide.
XZ
XZ chi-conotoxin; chi-MrIA; cone snail; inhibitor; amine transporter;
KW neuronal; noradrenaline transporter; urinary tract; analgesic; cardiant;
KW antiarrhythmic; antidepressant; anxiolytic; anti-inflammatory.
XZ
OS Conus marmoreus.
XZ
XZ WO200020444-A1.
XZ
PD 13-APR-2000.
XZ
XZ 01-OCT-1999; 99WO-AU00844.
XZ
XZ 02-OCT-1998; 98AU-0006274.
XZ (UYQU) UNIV QUEENSLAND.
XZ
XZ Lewis RJ, Alewood PF, Sharpe IA;
XZ WPI; 2000-303738/26.
XZ N-PSDB; AAA09112.
XZ
XZ Isolated, synthetic or recombinant chi-conotoxin peptide capable of
XZ inhibiting neuronal amine transporter used for treatment or prophylaxis
XZ of urinary or cardiovascular conditions, mood disorders, or
XZ treatment/control of pain/inflammation
XZ
XZ Example 7; Page 31; 47pp; English.
XZ
XZ This sequence is the conotoxin, chi-MrIA, a member of a new class of
XZ conotoxins, designated chi-conotoxin. It was isolated from the venom of
XZ the mollusc hunting cone snail, Conus marmoreus. The peptide is an
XZ inhibitor of the neuronal amine transporters, especially the neuronal
XZ noradrenaline transporter. Inhibitors of noradrenaline re-uptake which
XZ have a negligible anti-cholinergic effect are particularly useful in the
XZ treatment of lower urinary tract disorders. Chi-MrIA (0.1 nM-1 micro M)
XZ inhibited the accumulation of radiolabeled noradrenaline in a
XZ concentration-dependent manner, with a log IC-50 value of -8.17 plus or
XZ minus 0.0275 (n = 4). The concentration of chi-MrIA required to inhibit
XZ the accumulation by 50 percent was found to be approximately 7 nM. This
XZ concentration is approximately one order of magnitude lower than that
XZ needed for desipramine to produce the same effect. The peptides are
XZ useful for the treatment or prophylaxis of urinary or cardiovascular
XZ conditions or diseases (arrhythmia or coronary heart failure) or mood
XZ disorders (depression, anxiety or cravings), or the treatment or control
XZ of pain or inflammation (chronic pain, neuropathic pain or inflammatory
XZ pain).
XZ
SQ Sequence 61 AA;

Query Match 90.4%; Score 66; DB 21; Length 61;
Best Local Similarity 76.9%; Pred. No. 0.048;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NGVCCGXXLCHXC 13
DB 49 ngvcgkklchpc 61

RESULT 6
ID AAB08014
XZ AAB08014 standard; peptide; 12 AA.
XZ
AC AAB08014;
XZ
XZ 14-NOV-2000 (first entry)
XZ
XZ Generic formula for conotoxin peptide Mar2.
XZ
XZ Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar2.
XZ
XZ

```
OS Conus marmoreus.
FH Key Location/Qualifiers
FT Misc-difference 6 /label= Tyr, Xaa
FT /note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
FT O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT Misc-difference 7 /label= Lys, Xaa
FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT N,N,N-trimethyl Lys"
FT Misc-difference 11 /note= "optionally hydroxy-Pro"
FT WO200044769-A1.
PN
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-US01978.
XX
XX 29-JAN-1999; 99US-0118381.
XX 28-DEC-1999; 99US-0173343.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
XX
XX McIntosh JM, Olivera BM, Cruz LJ;
XX WPI; 2000-476222/41.
XX
XX Purified ap-conotoxin derived from cone snail venom for use as an
XX analgesic .
XX
XX Claim 2; Page 19; 29pp; English.
XX
XX The present sequence represents an ap-conotoxin peptide, designated
XX Mar2. Conotoxins are naturally available in minute amounts in the
XX venom of cone snails. The peptides have analgesic activity. The
XX peptides are used to treat or prevent pain.
XX
XX Sequence 12 AA;
SQ

Query Match 82.2%; Score 60; DB 21; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.078;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVC CGXXLCHXC 13
DB 1 gvc c g x x l c h p c 12

RESULT 7
AAB08015
ID AAB08015 standard; peptide; 12 AA.
XX
XX AAB08015;
XX
XX 14-NOV-2000 (first entry)
XX
XX Generic formula for conotoxin peptide U036.
XX
XX Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; U036.
XX
XX Conus marmoreus.
XX
XX Key Location/Qualifiers
FT Misc-difference 6 /label= Tyr, Xaa
FT /note= "Xaa is mono-halo-Tyr, di-halo-Tyr,
FT O-sulfo-Tyr, O-phospho Tyr or nitro-Tyr"
FT Misc-difference 7 /label= Lys, Xaa
FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
FT N,N,N-trimethyl Lys"
FT
FT
```

```
FT Misc-difference 11 /note= "optionally hydroxy-Pro"
FT N,N,N-trimethyl Lys"
XX
XX WO200044769-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-US01978.
XX
XX 29-JAN-1999; 99US-0118381.
XX 28-DEC-1999; 99US-0173343.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
XX
XX McIntosh JM, Olivera BM, Cruz LJ;
XX WPI; 2000-476222/41.
XX
XX Purified ap-conotoxin derived from cone snail venom for use as an
XX analgesic .
XX
XX Claim 2; Page 19; 29pp; English.
XX
XX The present sequence represents an ap-conotoxin peptide, designated
XX U036. Conotoxins are naturally available in minute amounts in the
XX venom of cone snails. The peptides have analgesic activity. The
XX peptides are used to treat or prevent pain.
XX
XX Sequence 12 AA;
SQ

Query Match 82.2%; Score 60; DB 21; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.078;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVC CGXXLCHXC 13
DB 1 gvc c g x x l c h p c 12

RESULT 8
AAB08018
ID AAB08018 standard; peptide; 12 AA.
XX
XX AAB08018;
XX
XX 14-NOV-2000 (first entry)
XX
XX Amino acid sequence of the conotoxin peptide Mar2.
XX
XX Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; Mar2.
XX
XX Conus marmoreus.
XX
XX Key Location/Qualifiers
FT Modified-site 12 /note= "hydroxy-Pro"
FT
XX
XX WO200044769-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-US01978.
XX
XX 29-JAN-1999; 99US-0118381.
XX 28-DEC-1999; 99US-0173343.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
XX
XX McIntosh JM, Olivera BM, Cruz LJ;
XX WPI; 2000-476222/41.
XX
```

XX PT Purified ap-conotoxin derived from cone snail venom for use as an
 PT analgesic -
 XX
 PS Claim 14; Page -: 29pp; English.
 XX
 CC The present sequence represents an ap-conotoxin peptide, designated
 CC Mar2. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.
 XX
 SQ Sequence 12 AA;
 Query Match 82.2%; Score 60; DB 21; Length 12;
 Best Local Similarity 75.0%; Pred. No. 0.078; 3; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 GVC CGXXLCHXC 13
 I I I I I I I I I
 Db 1 gvc cgyklchpc 12
 RESULT 9
 AAB08019
 ID AAB08019 standard; peptide; 12 AA.
 AC AAB08019;
 XX
 XX
 DT 14-NOV-2000 (first entry)
 DE Amino acid sequence of the conotoxin peptide U036.
 XX
 KW Ap-conotoxin; conotoxin; venom; cone snail; analgesic; pain; U036.
 XX
 OS Conus marmoreus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 7 /label= Lys, Xaa
 FT /note= "Xaa is N-methyl Lys, N,N-dimethyl Lys,
 FT N,N,N-trimethyl Lys"
 FT Modified-site 11
 FT /note= "hydroxy-Pro"
 XX
 PN WO200044769-A1.
 XX
 PD 03-AUG-2000.
 XX
 XX 28-JAN-2000; 2000WO-US01978.
 XX
 PR 29-JAN-1999; 99US-0118381.
 PR 28-DEC-1999; 99US-0173343.
 XX
 XX (UTAH) UNIV UTAH RES FOUND.
 XX
 PI McIntosh JM, Olivera BM, Cruz LJ;
 XX
 DR WPI; 2000-476222/41.
 XX
 PT Purified ap-conotoxin derived from cone snail venom for use as an
 PT analgesic -
 XX
 PS Claim 15; Page -: 29pp; English.
 XX
 CC The present sequence represents an ap-conotoxin peptide, designated
 CC U036. Conotoxins are naturally available in minute amounts in the
 CC venom of cone snails. The peptides have analgesic activity. The
 CC peptides are used to treat or prevent pain.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.

XX SQ Sequence 12 AA;
 Query Match 82.2%; Score 60; DB 21; Length 12;
 Best Local Similarity 83.3%; Pred. No. 0.078; 2; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 GVC CGXXLCHXC 13
 I I I I I I I I I
 Db 1 gvc cgyklchpc 12
 RESULT 10
 AAY92230
 ID AAY92230 standard; peptide; 13 AA.
 XX
 AC AAY92230;
 XX
 DT 10-AUG-2000 (first entry)
 DE Chi-conotoxin peptide, chi-MrIB.
 XX
 KW chi-conotoxin; chi-MrIB; cone snail; inhibitor; amine transporter;
 KW neuronal; noradrenaline transporter; urinary tract disorder; analgesic;
 KW antiarrhythmic; cardiac; antidepressant; anxiolytic; anti-inflammatory.
 XX
 OS Conus marmoreus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 12 /label= 4Hyp
 FT /note= "4-hydroxyproline"
 FT
 XX
 PN WO200020444-A1.
 XX
 PD 13-APR-2000.
 XX
 XX 01-OCT-1999; 99WO-AU00844.
 XX
 PR 02-OCT-1998; 98AU-0006274.
 XX
 PA (UYOU) UNIV QUEENSLAND.
 XX
 PI Lewis RJ, Alewood PF, Sharpe IA;
 XX
 DR WPI; 2000-303738/26.
 XX
 PT Isolated, synthetic or recombinant chi-conotoxin peptide capable of
 PT inhibiting neuronal amine transporter used for treatment or prophylaxis
 PT of urinary or cardiovascular conditions, mood disorders, or
 PT treatment/control of pain/inflammation
 XX
 PS Claim 3; Page 33; 47pp; English.
 XX
 CC This conotoxin, chi-MrIB, is a member of a new class of conotoxins,
 CC designated chi-conotoxin. It was isolated from the venom of the mollusc
 CC hunting cone snail, Conus marmoreus. The peptide is an inhibitor of the
 CC neuronal amine transporters, especially the neuronal noradrenaline
 CC transporter. Inhibitors of noradrenaline re-uptake which have a
 CC negligible anti-cholinergic effect are particularly useful in the
 CC treatment of lower urinary tract disorders. Chi-MrIA (0.1 nM-1 micro M)
 CC inhibited the accumulation of radiolabeled noradrenaline in a
 CC concentration-dependent manner, with a log IC-50 value of -8.17 plus or
 CC minus 0.0275 (n = 4). The concentration of chi-MrIA required to inhibit
 CC the accumulation by 50 percent was found to be approximately 7 nM. This
 CC concentration is approximately one order of magnitude lower than that
 CC needed for desipramine to produce the same effect. The peptides are
 CC useful for the treatment or prophylaxis of urinary or cardiovascular
 CC conditions or diseases (arrhythmia or coronary heart failure) or mood
 CC disorders (depression, anxiety or cravings), or the treatment or control
 CC of pain or inflammation (chronic pain, neuropathic pain or inflammatory
 CC pain).

```
SQ Sequence 13 AA;
Query Match 82.2%; Score 60; DB 21; Length 13;
Best Local Similarity 75.0%; Pred. No. 0.083;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVCCGXXLCHXC 13
   ||||| ||| |
Db 2 gvcggyklchpc 13

RESULT 11
AAU06111
ID AAU06111 standard; Protein; 538 AA.
XX
AC AAU06111;
XX
DT 24-OCT-2001 (first entry)
DE
DE Novel human polypeptide PROT1.
KW Human; PROT1; AJUBA-like; central nervous system disorder; CNS;
KW cancer; neuromuscular disorder; cardiac disorder; clone AL132780A.
XX
OS Homo sapiens.
XX
PN WO200158946-A2.
XX
PD 16-AUG-2001.
XX
PF 08-FEB-2001; 2001WO-US04402.
XX
PR 08-FEB-2000; 2000US-0180880.
PR 08-FEB-2000; 2000US-0181044.
PR 10-FEB-2000; 2000US-0181656.
PR 15-FEB-2000; 2000US-0182795.
PR 07-FEB-2001; 2001US-0182795.
XX
PA (CURA-) CURAGEN CORP.
PI
PI Taupier RJ, Majumder K, Vernet CAM, Prayaga SK;
XX
XX WPI: 2001-488970/53.
XX
XX N-PSDB; AAS09145.
XX
XX Isolated AJUBA-like, keratin-like and endothelin polypeptides for the
XX treatment and diagnosis of cancers and neurological disorders -
XX
XX Claim 1; Page 9-10; 141pp; English.
XX
XX The present invention relates to the isolation of 4 novel human
XX polypeptides, termed PROT1 polypeptides (AAU06111-AAU06114). The
XX PROT1 polypeptides are AJUBA-like, keratin-like or endothelin
XX polypeptides. Polynucleotides encoding PROT1, PROT2 polypeptides and
XX antibodies that bind PROT1 polypeptides are useful for treating or
XX preventing a pathology associated with altered levels of PROT1 especially
XX in a human subject. PROT1 polypeptides, PROT1 polynucleotides and PROT1
XX antibodies are useful for identifying an agent that binds to PROT1, for
XX identifying potential therapeutic agents for use in a treatment of a
XX pathology related to aberrant expression or physiological interaction
XX with PROT1, for screening for a modulator of activity of PROT1 and
XX for determining the presence of or a predisposition to a disease
XX associated with altered levels of PROT1. PROT1, an AJUBA-like protein,
XX is useful for treating central nervous system disorders, cancer,
XX Parkinson's disease, Alzheimer's disease, neuromuscular and cardiac
XX disorders. PROT2, a keratin, is useful for treating inflammatory and
XX neoplastic pancreatic disorders. PROT4, an endothelin, is useful for
XX treating hypertension, cardiovascular and neurological disorders. The
XX present sequence represents human polypeptide PROT1. The DNA sequence
XX encoding PROT1 is isolated from the genomic clone AL132780_A.
```

```
SQ Sequence 538 AA;
Query Match 60.3%; Score 44; DB 22; Length 538;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 4; Gaps 1;

QY 2 GVCC-----GXXLCHXC 13
   | | | | | | | | | |
Db 508 gcccfpldghllchgc 523

RESULT 12
AAU31406
ID AAU31406 standard; Protein; 93 AA.
XX
AC AAU31406;
XX
DT 18-DEC-2001 (first entry)
DE
DE Novel human secreted protein #1897.
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy -
XX
XX Claim 20; Page 432; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising
XX the nucleic acids encoding the polypeptides and cells genetically
XX engineered to express them are also useful for producing the proteins.
XX The proteins are useful in genetic vaccination, testing and
XX therapy, and can be used as nutritional supplements. They may be used to
XX increase stem cell proliferation; to regulate haematopoiesis; and in
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX immune suppression and/or stimulation; as anti-inflammatory agents; and
XX in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
XX sequences of novel human secreted proteins of the invention.
XX
XX Sequence 93 AA;
Query Match 58.9%; Score 43; DB 22; Length 93;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

QY 2 GVC CGXXLCH 11
| | | | |
Db 65 gpcgiffch 74

RESULT 13
AAM79259
ID AAM79259 standard; Protein; 766 AA.

XX AAM79259;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human protein SEQ ID NO 1921.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX 20-JUN-2000; 2000US-0598075.
XX 19-JUL-2000; 2000US-0620325.
XX 01-SEP-2000; 2000US-0654936.
XX *15-SEP-2000; 2000US-0663561.
XX 20-OCT-2000; 2000US-0693325.
XX 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.
XX N-PSDB; AAK52392.

XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -

XX Claim 20; Page 4320-4322; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78123-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 766 AA;

Query Match 58.9%; Score 43; DB 22; Length 766;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
| | | | |
Db 240 ccgnilch 247

RESULT 14
ABG05471
ID ABG05471 standard; Protein; 1203 AA.

XX ABG05471;
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #5462.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.
XX WO200175067-A2.
XX
XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS69658.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID NO 35830; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1203 AA;

Query Match 58.9%; Score 43; DB 22; Length 1203;
Best Local Similarity 75.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11

Db 737 ccgnilch 744

Db 1131 ccgnilch 1138

RESULT 15
ABG03989
ID ABG03989 standard; Protein; 1597 AA.

Search completed: May 19, 2002, 04:43:17
Job time: 12544 sec

XX AC ABG03989;

XX 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #3980.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS68176.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity

XX PS Claim 20; SEQ ID No 34348; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1597 AA;

Query Match 58.9%; Score 43; DB 22; Length 1597;
Best Local Similarity 75.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCGXXLCH 11
||| |||

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 01:52:58 ; Search time 60.36 seconds
(without alignments)
22.287 Million cell updates/sec

Title: US-09-580-201A-1

Perfect score: 51

Sequence: 1 XXXCCGXXXCX 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	78.4	400	1 ZBBE14	44.1K zinc-binding
2	39	76.5	286	2 S61199	hypothetical prote
3	39	76.5	1877	2 T21861	hypothetical prote
4	38	74.5	187	2 T10073	amine dehydrogenas
5	38	74.5	258	2 T30368	probable immediate
6	38	74.5	328	2 E97129	probable flavodoxi
7	38	74.5	948	2 A57640	retinoblastoma bin
8	38	74.5	1560	2 T42727	proliferation pote
9	37	72.5	186	2 A45910	ultra-high-sulfur
10	37	72.5	196	2 T25136	hypothetical prote
11	37	72.5	264	2 F97086	anaerobic sulfite
12	37	72.5	312	2 T25135	hypothetical prote
13	37	72.5	550	2 G85436	hypothetical prote
14	37	72.5	693	2 S49228	sodium-dependent p
15	36	70.6	404	2 T02396	hypothetical prote
16	36	70.6	514	2 T10359	hypothetical prote
17	36	70.6	670	2 T49510	hypothetical prote
18	36	70.6	756	2 D96527	fibroin-3 related
19	35	68.6	55	2 S25774	protein F27J15.24
20	35	68.6	127	2 T34270	testis-specific pr
21	35	68.6	428	2 T46025	hypothetical prote
22	34	66.7	215	2 T16342	hypothetical prote
23	33	64.7	412	2 F86296	hypothetical prote
24	32	62.7	64	2 B37915	hypothetical prote
25	32	62.7	141	2 S04925	ferredoxin [3Fe-4S
26	32	62.7	194	2 T25548	CTLA-2-beta protei
27	32	62.7	406	2 T04817	hypothetical prote
28	32	62.7	422	2 D86446	hypothetical prote
29	32	62.7	547	2 T30269	hypothetical prote

ALIGNMENTS

RESULT 1

ZBBE14

44.1K zinc-binding protein - ictalurid herpesvirus 1 (strain auburn 1)
C:Species: ictalurid herpesvirus 1

A:Note: host Ictalurus punctatus (channel catfish)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: D36794

R:Davidson, A.J.

submitted to GenBank, January 1992

A:Description: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A36804

A:Accession: D36794

A:Molecule type: DNA

A:Residues: 1-400 <DR>

A:Cross-references: GB:W75136; NID:g331209; PIDN:AAA88180.1; PID:g331286

R:Davidson, A.J.

Virology 186, 9-14, 1992

A:Title: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A39447; MUID:92087490

A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given

C:Genetics:

A:Gene: 78

C:Superfamily: ictalurid herpesvirus 44.1K zinc binding protein

C:Keywords: zinc finger

Query Match 78.4%; Score 40; DB 1; Length 400;
Best Local Similarity 50.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXXXC 13

Db 133 CCGATLCDSC 142

RESULT 2

S61199

hypothetical protein YDR313c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D9740.20

C:Species: Saccharomyces cerevisiae

C:Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 23-Mar-2001

C:Accession: S61199

R:Ding, H.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of S. cerevisiae cosmid 9740.

A:Reference number: S61160

A:Accession: S61199

A:Molecule type: DNA

A:Residues: 1-286 <DN>

A:Cross-references: EMBL:U28374; NID:g849207; PID:g849227; GSPDB:GN00004; MIPS:YDR31

C:Genetics:

A;Gene: MIPS:YDR313c
A;Map position: 4R

Query Match 76.5%; Score 39; DB 2; Length 286;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 41 CCGRIFCSC 50

RESULT 3

T21861

hypothetical protein F36F2.3 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T21861

R;Cottage, A.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19479

A;Accession: T21861

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1877 <WIL>

A;Cross-references: EMBL:Z81532; PIDN:CAB04326.1; GSPDB:GN00019; CESP:F36F2.3

A;Experimental source: clone F36F2

C;Genetics:

A;Gene: CRSP:F36F2.3

A;Map position: 1

A;Introns: 49/3; 86/2; 112/3; 139/2; 235/3; 284/3; 436/3; 507/2; 566/2; 678/2; 1000/3; 1

Query Match

Best Local Similarity 76.5%; Score 39; DB 2; Length 1877;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 320 CCGNSYCADC 329

RESULT 4

T10073

amine dehydrogenase (EC 1.4.99.3) light chain - *Methylophilus methylotrophus*

C;Species: *Methylophilus methylotrophus*

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000

C;Accession: T10073

R;Chistoserdov, A.Y.; McIntire, W.S.; Mathews, F.S.; Lidstrom, M.E.

J. Bacteriol. 176, 4073-4080, 1994

A;Title: Organization of the methylamine utilization (mau) genes in *Methylophilus methylotrophus*

A;Reference number: Z16936; MUID:94292427

A;Accession: T10073

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-187 <CHI>

A;Cross-references: EMBL:L26407; NID:g561931; PIDN:AAB46951.1; PID:g556334

A;Experimental source: strain W3A1

C;Genetics:

A;Gene: mauA

C;Complex: heterotetramer; two large and two small chains

C;Superfamily: amine dehydrogenase light chain

C;Keywords: oxidoreductase

Query Match

Best Local Similarity 74.5%; Score 38; DB 2; Length 187;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 133 CCGKQTCRC 142

RESULT 5

T30368

probable immediate-early transactivator 0 - *Lymantria dispar* nuclear polyhedrosis vir

C;Species: *Lymantria dispar* nuclear polyhedrosis virus, LdMNPV

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C;Accession: T30368

R;Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; R

Virol. 253, 17-34, 1999

A;Title: Sequence and analysis of the genome of a baculovirus pathogenic for *Lymantria*

A;Reference number: Z20836; MUID:99124785

A;Accession: T30368

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-258 <KUZ>

A;Cross-references: EMBL:AF081810; PIDN:AAC70206.1

C;Keywords: immediate-early protein

Query Match

Best Local Similarity 74.5%; Score 38; DB 2; Length 258;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 209 CCGYRVCNAC 218

RESULT 6

E97129

probable flavodoxin oxidoreductase CAC1860 [imported] - *Clostridium acetobutylicum*

C;Species: *Clostridium acetobutylicum*

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C;Accession: E97129

R;Nolling, J.; Brennon, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: E97129

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-328 <KUR>

A;Cross-references: GB:AE001437; PIDN:AAK79824.1; PID:gl5024837; GSPDB:GN00168

A;Experimental source: *Clostridium acetobutylicum* ATCC824

C;Genetics:

A;Gene: CAC1860

Query Match

Best Local Similarity 74.5%; Score 38; DB 2; Length 328;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 290 CCGEGICGAC 299

RESULT 7

A57640

retinoblastoma binding protein RBQ-1 - human

C;Species: *Homo sapiens* (man)

C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 17-Mar-2000

C;Accession: A57640

R;Sakai, Y.; Saijo, M.; Coelho, K.; Kishino, T.; Nilkawa, N.; Taya, Y.

Genomics 30, 98-101, 1995

A;Title: cDNA sequence and chromosomal localization of a novel human protein, RBQ-1

A;Reference number: A57640; MUID:96129310

A;Accession: A57640

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-948 <SAK>

A:Cross-references: GB:X85133; NID:g728590; PIDN:CAA59445.1; PID:g755748
C:Genetics:
A:Gene: GDB:RBBP6
A:Cross-references: GDB:626076
A:Map position: 16p12-16p11.2
C:Superfamily: RING finger homology
C:Keywords: tandem repeat; zinc
F:79-129/Domain: RING finger homology <RNG>

Query Match 74.5%; Score 38; DB 2; Length 948;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 98 CCGNSYCDEC 107

RESULT 8
T42727
proliferation potential-related protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C:Accession: T42727
R:Witte, M.M.; Scott, R.E.
submitted to the EMBL Data Library, November 1998
A:Reference number: 22246
A:Accession: T42727
A:Status: Preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1560 <WIT>
A:Cross-references: EMBL:U83913; NID:g3858884; PID:g3858885; PIDN:AAC72432.1
A:Experimental source: strain Balb/C
C:Genetics:
A:Gene: P2P-R
C:Function:
A:Description: involved in hnRNP association and Rb1 binding
C:Superfamily: RING finger homology
F:57-107/Domain: RING finger homology <RRN>

Query Match 74.5%; Score 38; DB 2; Length 1560;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 76 CCGNSYCDEC 85

RESULT 9
A45910
ultra-high-sulfur keratin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 24-Sep-1999
C:Accession: A45910
R:McNab, A.R.; Wood, L.; Theriault, N.; Gierman, T.; Vogeli, G.
J. Invest. Dermatol. 92, 263-266, 1989
A:Title: An ultra-high sulfur keratin gene is expressed specifically during hair growth.
A:Reference number: A45910; MUID:89140394
A:Accession: A45910
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <MCN>
A:Cross-references: GB:M27685; NID:g341749; PIDN:AAA81560.1; PID:g1066818
C:Superfamily: ultra-high-sulfur keratin

Query Match 72.5%; Score 37; DB 2; Length 186;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 42 CCGSSCCQPC 51

RESULT 10
T25136
hypothetical protein T22H2.5a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25136
R:Lennard, N.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19985
A:Accession: T25136
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-196 <WIL>
A:Cross-references: EMBL:Z81595; PIDN:CAB04755.1; GSPDB:GN00019; CESP:T22H2.5a
A:Experimental source: clone T22H2
C:Genetics:
A:Gene: CESP:T22H2.5a
A:Map position: 1
A:Introns: 25/3; 126/3; 161/1; 180/3; 196/1

Query Match 72.5%; Score 37; DB 2; Length 196;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 58 CCGCLACIGC 67

RESULT 11
F97086
anaerobic sulfite reductase, B chain [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: F97086
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: F97086
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79481.1; PID:g15024461; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1514

Query Match 72.5%; Score 37; DB 2; Length 264;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 231 CCGIGKCGHC 240

RESULT 12
T25135
hypothetical protein T22H2.5b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T25135
R:Lennard, N.
submitted to the EMBL Data Library, November 1996

A:Reference number: Z19985
A:Accession: T25135
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-312 <MTL>
A:Cross-references: EMBL:Z81595; PIDN:CAB04754.1; GSPDB:GN00019; CESP:T22H2.5b
A:Experimental source: clone T22H2
C:Genetics:
A:Gene: CESP:T22H2.5b
A:Map position: 1
A:Introns: 135/3; 236/3; 271/1; 290/3; 306/1
C:Superfamily: Caenorhabditis elegans hypothetical protein K08D10.7

Query Match 72.5%; Score 37; DB 2; Length 312;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 168 CCGCLACIGC 177

RESULT 13
G85436
hypothetical protein AT4g36980 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: G85436
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488
A:Accession: G85436
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-550 <STO>
A:Cross-references: GB:NC_001268; NID:g7270647; PIDN:CAB80364.1; GSPDB:GN00140
A:Gene: AT4g36980
A:Map position: 4

Query Match 72.5%; Score 37; DB 2; Length 550;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 499 CCGSFFCCLC 508

RESULT 14
S49228
sodium-dependent phosphate transporter - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 05-Nov-1999
C:Accession: S68972; S49228
R:Helps, C.; Murer, H.; McGivan, J.
Eur. J. Biochem. 228, 927-930, 1995
A:Title: Cloning, sequence analysis and expression of the cDNA encoding a sodium-depende
A:Reference number: S68972; MUID:95255303
A:Accession: S68972
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-693 <HE2>
A:Cross-references: EMBL:X81699; NID:g547483; PIDN:CAA57345.1; PID:g547484

Query Match 72.5%; Score 37; DB 2; Length 693;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 628 CCGLCGCSKC 637

RESULT 15
T02396
hypothetical protein At2g44410 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F411.22
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 03-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02396; B84878
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; K.
submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC F411 genomic sequence.
A:Reference number: Z14667
A:Accession: T02396
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-404 <ROU>
A:Cross-references: EMBL:AC004521; NID:g3128166; PID:g3128183
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: B84878
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-404 <STO>
A:Cross-references: GB:AE002093; NID:g3128183; PIDN:AAC16087.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g44410; F411.22
A:Map position: 2
C:Superfamily: RING finger homology
F:112-162/Domain: RING finger homology <RRN>

Query Match 70.6%; Score 36; DB 2; Length 404;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
DB 130 CCGHLFCWGC 139

Search completed: May 19, 2002, 04:48:46
Job time: 10548 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 04:44:14 ; Search time 42.56 Seconds
(without alignments)
12.737 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCGXXXCX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	40	78.4	400	1	VG78_HSV11	Q00167 ictaluriid h
2	38	74.5	186	1	DHML_METFL	Q50425 methylobaci
3	38	74.5	187	1	DHML_METME	Q59543 methylobaci
4	35	68.6	55	1	M84C_DROME	Q01644 drosophila
5	33	64.7	266	1	GNP1_GIALA	O97439 giardia lam
6	32	62.7	64	1	FER2_STRGO	P18325 streptomyc
7	32	62.7	75	1	MT_EISEFO	P81695 eisenia foe
8	32	62.7	141	1	CT2B_MOUSE	P12400 mus musculu
9	32	62.7	1239	1	POLS_EEEV	P08768 eastern equ
10	32	62.7	1240	1	POLS_EEEV3	P27284 eastern equ
11	32	62.7	1247	1	POLS_ONNVG	P22056 o'nyong-nyo
12	32	62.7	1253	1	POLS_SFV	P03315 semliki for
13	32	62.7	1254	1	POLS_RRVN	P13890 ross river
14	32	62.7	1254	1	POLS_RRVT	P08491 ross river
15	31	60.8	65	1	TXM7_DENAN	P80970 dendroaspis
16	31	60.8	72	1	YXWU_VACCC	P20530 vaccinia vi
17	31	60.8	86	1	TXM2_DENAN	P18328 dendroaspis
18	31	60.8	99	1	NIC1_HUMAN	O9uq19 homo sapien
19	31	60.8	155	1	NEU4_CATCO	P16229 catostomus
20	31	60.8	198	1	DDL_STRMU	P95803 streptococc
21	31	60.8	233	1	TPL_TREPA	P45685 treponema p
22	31	60.8	239	1	NOO2_PARDE	P29914 paracoccu
23	31	60.8	245	1	IEO_NPVOP	O10369 orgyia pseu
24	31	60.8	261	1	IEO_NPVAC	P41710 autographa
25	31	60.8	299	1	ALC_RABIT	P01879 oryctolagus
26	31	60.8	352	1	DMPP_PSESP	P19734 pseudomonas
27	31	60.8	380	1	APJ_HUMAN	P35414 homo sapien
28	31	60.8	380	1	APJ_MACMU	O97666 macaca mula
29	31	60.8	415	1	YLN2_CAEEL	O18964 caenorhabdi
30	31	60.8	458	1	PRTC_RABIT	Q28661 oryctolagus
31	31	60.8	459	1	PRTC_PIG	O9gip2 sus scrofa
32	31	60.8	461	1	PRTC_HUMAN	P04070 homo sapien
33	31	60.8	461	1	PRTC_MOUSE	P33587 mus musculu

34	31	60.8	461	1	PRTC_RAT	P31394 rattus norv
35	31	60.8	885	1	TRP2_RAT	O9r283 rattus norv
36	31	60.8	1172	1	TRP2_MOUSE	O9r244 mus musculu
37	31	60.8	1173	1	VGL2_CVH22	P15423 human coron
38	31	60.8	1236	1	POLS_WEEV	P13897 western equ
39	31	60.8	1245	1	POLS_SINDO	P27285 sindbis vir
40	31	60.8	1245	1	POLS_SINDV	P03316 sindbis vir
41	31	60.8	1254	1	POLS_EEVV8	P05674 venezuelan
42	31	60.8	1254	1	POLS_EEVVE	P36330 venezuelan
43	31	60.8	1254	1	POLS_EEVVM	P36331 venezuelan
44	31	60.8	1254	1	POLS_EEVT	P09592 venezuelan
45	31	60.8	1255	1	POLS_EEVV3	P36329 venezuelan

ALIGNMENTS

RESULT 1
VG78_HSV11
ID VG78_HSV11 STANDARD; PRT; 400 AA.
AC Q00167;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical gene 78 zinc-binding protein.
GN 78.
OS Ictaluriid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC unclassified Herpesviridae.
OX NCBI_TaxID=10401;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AUBURN 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davidson A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M751136; AAA88180.1; -
CC PIR; D36794; ZBBE14.
KW Hypothetical protein; Zinc; Zinc-finger.
SQ SEQUENCE 400 AA; 44167 MW; D941DB7738B95CA9 CRC64;

Query Match 78.4%; Score 40; DB 1; Length 400;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCX 13
Db 133 CCGATLCDSC 142

RESULT 2
DHML_METFL STANDARD; PRT; 186 AA.
ID DHML_METFL
AC Q50425;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Methylamine dehydrogenase light chain precursor (EC 1.4.99.3) (MADH).
GN MAUA.
OS Methylobacillus flagellatum.
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
OC Methylobacillus.

```

OX NCBI_TaxID=405;
RP [1]
RN SEQUENCE FROM N.A.
RC STRAIN=KT / ATCC 51484 / DSM 6875 / VKM B-1610;
RX MEDLINE=95362696; PubMed=7635847;
RA Gak E.R., Chistoserdov A.Y., Lidstrom M.E.;
RT "Cloning, sequencing, and mutation of a gene for azurin in
RT Methylobacillus flagellatum KT.";
RL J. Bacteriol. 177:4575-4578(1995).
CC -1- FUNCTION: METHYLAMINE DEHYDROGENASE CARRIES OUT THE OXIDATION OF
CC METHYLAMINE. ELECTRONS ARE PASSED FROM METHYLAMINE DEHYDROGENASE
CC TO AMICVANIN.
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + acceptor = RCHO + NH(3)
CC + reduced acceptor.
CC -1- COFACTOR: EACH OF THE SMALL SUBUNITS CONTAINS A COVALENTLY BOUND
CC PQQ-LIKE COFACTOR CALLED TRYPTOPHAN TRYPTOPHYLAQUINONE (TTQ).
CC -1- PATHWAY: METHYLAMINE UTILIZATION.
CC -1- SUBUNIT: TETRAMER OF TWO LIGHT AND TWO HEAVY CHAINS.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: HIGH, TO OTHER SPECIES MADH LIGHT CHAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L37427; AAC41474.1; -
DR HSP; P22619; 2BBK
DR InterPro; IPR004229; Me-amine-deh_L.
DR Pfam; PF02975; Me-amine-deh_L; 1.
KW Oxidoreductase; Electron transport; Periplasmic; TTQ; Signal.
FT SIGNAL 1 56 POTENTIAL.
FT CHAIN 57 186 METHYLAMINE DEHYDROGENASE LIGHT CHAIN.
FT BINDING 112 112 TTQ (BY SIMILARITY).
FT BINDING 163 163 TTQ (BY SIMILARITY).
FT DISULFID 78 143 BY SIMILARITY.
FT DISULFID 84 116 BY SIMILARITY.
FT DISULFID 91 176 BY SIMILARITY.
FT DISULFID 93 141 BY SIMILARITY.
FT DISULFID 101 132 BY SIMILARITY.
FT DISULFID 133 164 BY SIMILARITY.
SQ SEQUENCE 186 AA; 20174 MW; E94721FDA794B523 CRC64;

Query Match 74.5%; Score 38; DB 1; Length 186;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
Db 132 CCGKQTGRC 141
III I I

RESULT 3
DHML_METME STANDARD; PRT; 187 AA.
AC Q59543;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Methylamine dehydrogenase light chain precursor (EC 1.4.99.3) (MADH).
GN NAUA.
OS Methylophilus methylotrophus (Bacterium W3A1).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
OC Methylophilus.
OX NCBI_TaxID=17;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94292427; PubMed=8021188;
RA Chistoserdov A.Y., McIntire W.S., Mathews F.S., Lidstrom M.E.;

```

```

RT "Organization of the methylamine utilization (mau) genes in
RT Methylophilus methylotrophus W3A1-NS.";
RL J. Bacteriol. 176:4073-4080(1994).
CC -1- FUNCTION: METHYLAMINE DEHYDROGENASE CARRIES OUT THE OXIDATION OF
CC METHYLAMINE. ELECTRONS ARE PASSED FROM METHYLAMINE DEHYDROGENASE
CC TO AMICVANIN.
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + acceptor = RCHO + NH(3)
CC + reduced acceptor.
CC -1- COFACTOR: EACH OF THE SMALL SUBUNITS CONTAINS A COVALENTLY BOUND
CC PQQ-LIKE COFACTOR CALLED TRYPTOPHAN TRYPTOPHYLAQUINONE (TTQ).
CC -1- PATHWAY: METHYLAMINE UTILIZATION.
CC -1- SUBUNIT: TETRAMER OF TWO LIGHT AND TWO HEAVY CHAINS.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: HIGH, TO OTHER SPECIES MADH LIGHT CHAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L26407; AAB46951.1; -
DR HSP; P22619; 2BBK.
DR InterPro; IPR004229; Me-amine-deh_L.
DR Pfam; PF02975; Me-amine-deh_L; 1.
KW Oxidoreductase; Electron transport; Periplasmic; TTQ; Signal.
FT SIGNAL 1 57 POTENTIAL.
FT CHAIN 58 187 METHYLAMINE DEHYDROGENASE LIGHT CHAIN.
FT BINDING 113 113 TTQ (BY SIMILARITY).
FT BINDING 164 164 TTQ (BY SIMILARITY).
FT DISULFID 79 144 BY SIMILARITY.
FT DISULFID 92 177 BY SIMILARITY.
FT DISULFID 94 142 BY SIMILARITY.
FT DISULFID 102 133 BY SIMILARITY.
FT DISULFID 134 165 BY SIMILARITY.
SQ SEQUENCE 187 AA; 20237 MW; E33FEBAE30CC5CED CRC64;

Query Match 74.5%; Score 38; DB 1; Length 187;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
Db 133 CCGKQTGRC 142
III I I

RESULT 4
M84C_DROME STANDARD; PRT; 55 AA.
AC Q01644; Q9VIA0;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Male specific sperm protein Mst84Dc.
GN Mst84DC OR CG17945
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=92102953; PubMed=1684716;
RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,
RA Schaefer M.;
RT "A cluster of four genes selectively expressed in the male germ line
RT of Drosophila melanogaster.";
RL Mech. Dev. 35:143-151(1991).

```


[2]
SEQUENCE FROM N.A.
STRAIN-BERKELEY:
MEDLINE-20196006; PubMed-10731132;
ADAMS M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos B.L.G.,
Abrill J.F., Agayani A., An H.-J., Andrews-Franckoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brodsky P., Brothier P.,
Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J.S., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
"The genome sequence of *Drosophila melanogaster*."
Science 287:2185-2195(2000).
-!- TISSUE SPECIFICITY: TESTIS.
-!- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
-!- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
MOTIFS.
-!- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X67703; CAA47939.1; -
DB EMBL: AE003672; AAF54025.1; -
DR HSSP: P01180; INPO.
DR FlyBase: FBgn0004174; Mst84Bc.
KW Spermatogenesis; Repeat; Multigene family.
SQ SEQUENCE 55 AA; 5225 MW; 95A12F3AEC88BD6C CRC64;

Query Match 68.6%; Score 35; DB 1; Length 55;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 CCGXXXCCXC 13
||| |
DB 9 CCGYCCGCP 18

RESULT 5
GNP1_GIALA STANDARD; PRT; 266 AA.
AC 097439;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucosamine-6-phosphate isomerase 1 (EC 3.5.99.6) (Glucosamine-
6-phosphate deaminase 1) (GNPDA 1) (GlcN6P deaminase 1).
GN GP11
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WR4 / POLISH GENOTYPE;
RC van Keulen H., Steimle P.A., Bulik D.A., Borowiak R.K., Jarroll E.L.;
RT "Cloning of two Giardia glucosamine 6-phosphate isomerase genes only
one of which is transcriptionally activated during encystation.";
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RL -!- CATALYTIC ACTIVITY: D-glucosamine 6-phosphate + H(2)O = D-fructose
6-phosphate + NH(3).
CC -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF050754; AAD02508.1; -
DR HSSP: P09375; IDEA.
DR InterPro: IPR000457; Glucosamine_iso.
DR Pfam: PF01182; Glucosamine_iso; 1.
DR PROSITE: PS01161; GLC_GALNAC_ISOMERASE; 1.
KW Hydrolase; Carbohydrate metabolism.
FT ACT SITE 67 67 BY SIMILARITY
SQ SEQUENCE 266 AA; 29406 MW; 3EC301B8D868386B CRC64;
Query Match 64.7%; Score 33; DB 1; Length 266;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 CCGXXXC 10
||| |
DB 255 CCGSTSC 261
RESULT 6
FER2_STRGO STANDARD; PRT; 64 AA.
AC P18325;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ferredoxin 2 (Fd-2).
GN SUBB.
OS Streptomyces griseolus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1909;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-43.
RC STRAIN-ATCC 11796;
RX MEDLINE-91105127; PubMed-1846297;
RA O'Keefe D.P., Gibson K.J., Emptage M.H., Lenstra R., Romesser J.A.,
RA Little P.J., Omer C.A.;
RT "Ferredoxins from two sulfonyleurea herbicide monooxygenase systems in
Streptomyces griseolus.";

RL Biochemistry 30:447-455(1991).
CC -1- FUNCTION: ELECTRON TRANSPORT PROTEIN FOR THE CYTOCHROME P-450-SU2
CC SYSTEM.
CC -1- COFACTOR: BINDS A 3FE-4S CLUSTER.
CC -1- INDUCTION: By herbicides.
CC -1- SIMILARITY: >50%, TO FERREDOXIN-1, AND TO B-STEATOTHERMOPHILUS,
CC C-THERMOACETICUM, AND D.AFRICANUS FERREDOXINS.
CC -1- SIMILARITY: TO THE N-TERMINAL OF R.FASCIANS FERREDOXIN PAS2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M32239; AAA26826.1; -
CC PIR; B37915; B37915.
CC HSSP; P00209; IFXD.
CC InterPro; IPR001080; 3Fe4S-ferredoxin.
CC PRINTS; PR00352; 3FE4SFRDOXIN.
CC Electron transport; Iron-sulfur; 3Fe-4S.
CC METAL 10 10 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
CC METAL 16 16 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
CC METAL 55 55 IRON-SULFUR (3FE-4S) (BY SIMILARITY).
CC SEQUENCE 64 AA; 6647 MW; 3918DF9BDAEE6D0C CRC64;
CC -----
CC Query Match 62.7%; Score 32; DB 1; Length 64;
CC Best Local Similarity 57.1%; Pred. No. 94;
CC Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC QY 4 CCGXXXC 10
CC III I
CC Db - 10 CCGAGSC 16
CC -----
CC RESULT 7
CC MT_EISFO STANDARD; PRT; 75 AA.
CC AC P81695; O96822;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DE Cadmium-metallothionein (MT) (Fragment).
CC GN MT.
CC OS Eisenia foetida (Common branding worm) (Common dung-worm).
CC OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
CC OC Lumbricina; Lumbricidae; Eisenia.
CC OX NCBI_TaxID=6396;
CC RN [1]
CC SEQUENCE OF 1-41.
CC RA Gruber C., Dallinger R., Hunziker P., Gehrig P., Berger B.,
CC RA Stuerzenbaum S.;
CC RT "(Cd)-metallothionein from Eisenia foetida: evidence for
CC RT posttranslational processing to a functional one-domain protein.";
CC RL Submitted (FEB-1999) to the SWISS-PROT data bank.
CC RN [2]
CC SEQUENCE OF 2-75 FROM N.A.
CC RA Sturzenbaum S.R., Guppy L., Morgan A.J., Kille P.;
CC RT "Metallothionein isolated from the common branding worm Eisenia
CC RT foetida";
CC RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC CC -1- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC CC -1- TISSUE SPECIFICITY: POSTERIOR ALIMENTARY CANAL.
CC CC -1- INDUCTION: BY CADMIUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ236886; CAAL5423.1; -
CC InterPro; IPR001396; Metallothion_Echnd.
CC PRINTS; PR00873; MTECHNOIDEA.
CC KW Metal-binding; Metal-thiolate cluster; Chelation; Cadmium.
CC FT NON_TER 1 1 A -> O (IN REF. 1).
CC FT CONFLICT 11 11
CC SEQUENCE 75 AA; 7411 MW; FDE2C30BA606494B CRC64;
CC -----
CC Query Match 62.7%; Score 32; DB 1; Length 75;
CC Best Local Similarity 57.1%; Pred. No. 1e+02;
CC Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC QY 4 CCGXXXC 10
CC III I
CC Db 4 CCGKSTC 10
CC -----
CC RESULT 8
CC CT2B_MOUSE STANDARD; PRT; 141 AA.
CC AC P12400;
CC DT 01-OCT-1989 (Rel. 12, Created)
CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC DE CTLA-2-beta protein precursor (Fragment).
CC GN CTLA2B.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC RN [1]
CC SEQUENCE FROM N.A.
CC RP STRAIN=C57BL/10-BR; TISSUE=T-cell;
CC RX MEDLINE=89276474; PubMed=2786470;
CC RA Denizot F., Brunet J.-F., Roustan P., Harper K., Suzan M.,
CC RA Luciani M.-F., Mattei M.-G., Golstein P.;
CC RT "Novel structures CTLA-2 alpha and CTLA-2 beta expressed in mouse
CC RT activated T cells and mast cells and homologous to cysteine
CC RT proteinase proteoglycans.";
CC RL Eur. J. Immunol. 19:631-635(1989).
CC CC -1- FUNCTION: NOT KNOWN, EXPRESSED IN ACTIVATED T-CELL.
CC CC -1- SIMILARITY: STRONG, TO THE PROPEPTIDE REGIONS OF CYSTEINE
CC CC PROTEASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15592; CAA33615.1; -
CC PIR; S04925; S04925.
CC HSSP; P07711; LCS8.
CC MGD; MGI:88555; Ctla2b.
CC KW T-cell; Repeat; Signal.
CC FT NON_TER 1 1
CC FT SIGNAL <1 31
CC FT CHAIN 32 141
CC FT DOMAIN 43 48
CC FT REPEAT 43 45
CC FT REPEAT 46 48
CC SEQUENCE 141 AA; 16248 MW; 232E2EA32B8F2AC0 CRC64;
CC -----
CC POTENTIAL.
CC CTLA-2-BETA PROTEIN.
CC 2 X 3 AA TANDEM REPEATS OF E-W-K.

Query Match 62.7%; Score 32; DB 1; Length 141;
 Best Local Similarity 57.1%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGXXXC 10
 ||| |
 DB 107 CCGSSMC 113

RESULT 9

POLY_EEV3 STANDARD; PRT; 1239 AA.
 AC P08768;
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Structural polyprotein (P130) [Contains: Coat protein C (EC 3.4.21.-)
 DE (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
 DE 6 kDa peptide; Spike glycoprotein E1].
 OS Eastern equine encephalitis virus (Eastern equine encephalomyelitis
 OS virus).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 OX NCBI_TaxID=11021;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=82V-2137;
 RA Chang G.-J.J., Trent D.W.;
 RT "Nucleotide sequence of the genome region encoding the 26S mRNA of
 RT eastern equine encephalomyelitis virus and the deduced amino acid
 RT sequence of the viral structural proteins.";
 RL J. Gen. Virol. 68:2129-2142(1987).
 CC --!- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
 CC --!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC --!- MISCELLANEOUS: THE 6 kDa POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
 CC SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL
 CC HEMAGGLUTININ.
 CC --!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X05816; CAA29261.1; -
 CC PIR; A26816; VHWVEE.
 CC HSSP; P03315; 1VCP.
 CC MEROPS; S03.001; -
 CC InterPro; IPR002548; Alpha_E1_glycop.
 CC InterPro; IPR000936; Alpha_E2_glycop.
 CC InterPro; IPR002533; Alpha_E3_glycop.
 CC InterPro; IPR001836; Alpha_core.
 CC InterPro; IPR000930; Togavirin.
 CC Pfam; PF00944; Alpha_core; 1.
 CC Pfam; PF01589; Alpha_E1_glycop; 2.
 CC Pfam; PF00943; Alpha_E2_glycop; 1.
 CC Pfam; PF01563; Alpha_E3_glycop; 1.
 CC PRINTS; PR00798; TOGAVIRIN.
 CC Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;
 KW Serine protease.
 FT CHAIN 1 259 COAT PROTEIN C.
 FT CHAIN 260 322 SPIKE GLYCOPROTEIN E3.
 FT CHAIN 323 742 SPIKE GLYCOPROTEIN E2.
 FT CHAIN 743 798 6 KDA PEPTIDE.
 FT CHAIN 799 1239 SPIKE GLYCOPROTEIN E1.
 FT ACT_SITE 136 136 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 142 142 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 210 210 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT TRANSMEM 261 277 POTENTIAL.

FT TRANSMEM 684 701 POTENTIAL.
 FT TRANSMEM 727 737 POTENTIAL.
 FT TRANSMEM 777 798 POTENTIAL.
 FT TRANSMEM 1211 1235 POTENTIAL.
 FT CARBOHYD 49 49
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 932 932 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1239 AA; 137431 MW; 8C7664A405D2D41C CRC64;

Query Match 62.7%; Score 32; DB 1; Length 1239;
 Best Local Similarity 57.1%; Pred. No. 4.8e+02;
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGXXXC 10
 ||| |
 DB 860 CCGATQC 866

RESULT 10

POLY_EEV3 STANDARD; PRT; 1240 AA.
 AC P27284;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Structural polyprotein (P130) [Contains: Coat protein C (EC 3.4.21.-)
 DE (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
 DE 6 kDa peptide; Spike glycoprotein E1].
 OS Eastern equine encephalitis virus (strain va33(ten broeck)) (Eastern
 OS equine encephalomyelitis virus).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 OX NCBI_TaxID=11022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=91220727; PubMed=2024496;
 RA Weaver S.C., Scott T.W., Rico-Hesse R.;
 RT "Molecular evolution of eastern equine encephalomyelitis virus in
 RT North America.";
 RL Virology 182:774-784(1991).
 CC --!- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
 CC --!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC --!- MISCELLANEOUS: THE 6 kDa POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
 CC SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL
 CC HEMAGGLUTININ.
 CC --!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M69094; AAA42980.1; -
 CC PIR; A39992; VHWVEV.
 CC HSSP; P03315; 1VCP.
 CC MEROPS; S03.001; -
 CC InterPro; IPR002548; Alpha_E1_glycop.
 CC InterPro; IPR000936; Alpha_E2_glycop.
 CC InterPro; IPR002533; Alpha_E3_glycop.
 CC InterPro; IPR001836; Alpha_core.
 CC InterPro; IPR000930; Togavirin.
 CC Pfam; PF00944; Alpha_core; 1.
 CC Pfam; PF01589; Alpha_E1_glycop; 2.
 CC Pfam; PF00943; Alpha_E2_glycop; 1.
 CC Pfam; PF01563; Alpha_E3_glycop; 1.
 CC PRINTS; PR00798; TOGAVIRIN.
 CC Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;
 KW Serine protease.
 FT CHAIN 1 260 COAT PROTEIN C.
 FT TRANSMEM 1 260

```
FT CHAIN 261 323 SPIKE GLYCOPROTEIN E3.
FT CHAIN 324 743 SPIKE GLYCOPROTEIN E2.
FT CHAIN 744 799 6 KDA PEPTIDE.
FT CHAIN 800 1240 SPIKE GLYCOPROTEIN E1.
FT ACT_SITE 137 137 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 143 143 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 211 211 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT TRANSMEM 259 276 POTENTIAL.
FT TRANSMEM 695 712 POTENTIAL.
FT TRANSMEM 722 738 POTENTIAL.
FT TRANSMEM 781 799 POTENTIAL.
FT TRANSMEM 1212 1236 POTENTIAL.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 834 834 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 933 933 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1240 AA; 137290 MW; AEBEB1599D083045 CRC64;

Query Match 62.7%; Score 32; DB 1; Length 1240;
Best Local Similarity 57.1%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 CCGXXXC 10
    |||
Db 861 CCGATQC 867

RESULT 11
POL_S_ONVNG STANDARD; PRT; 1247 AA.
AC P22056;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Structural polyprotein (PI30) [Contains: Coat protein C (EC 3.4.21.-)
DE (capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
DE 6 kDa peptide; Spike glycoprotein E1].
OS O'nyong-nyong virus (strain Gulu) (ONN).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=11028;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90177206; PubMed=2155505;
RA Levinson R.S., Strauss J.H., Strauss E.G.;
RT "Complete sequence of the genomic RNA of O'nyong-nyong virus and its
RT use in the construction of alphavirus phylogenetic trees.";
RL Virology 175:110-123(1990).
CC -1- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THE 6 kDa POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
CC SEQUENCE FOR THE SPIKE GLYCOPROTEIN E1, WHICH IS THE VIRAL
CC HEMAGGLUTININ.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL: M20303; AAA46785.1; -
CC PIR: B34680; VHWVW2.
CC HSP: P03315; IVP.
CC MEROPS: S03.001; -
CC InterPro: IPR002548; Alpha_E1_glycop.
CC InterPro: IPR000936; Alpha_E2_glycop.
CC InterPro: IPR002533; Alpha_E3_glycop.
CC InterPro: IPR001836; Alpha_core.
CC InterPro: IPR001271; Defensin.
```

```
DR InterPro: IPR000930; Togavirin.
DR Pfam: PF00944; Alpha_core; 1.
DR Pfam: PF01589; Alpha_E1_glycop; 1.
DR Pfam: PF00943; Alpha_E2_glycop; 1.
DR Pfam: PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
DR SMART; SM00048; DEFSN; 1.
KW Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;
KW Serine protease.
FT CHAIN 1 260 COAT PROTEIN C.
FT CHAIN 261 324 SPIKE GLYCOPROTEIN E3.
FT CHAIN 325 747 SPIKE GLYCOPROTEIN E2.
FT CHAIN 748 808 6 KDA PEPTIDE.
FT CHAIN 809 1247 SPIKE GLYCOPROTEIN E1.
FT ACT_SITE 138 138 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 144 144 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 949 949 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1247 AA; 137969 MW; B7DCBDD17D9563CB CRC64;

Query Match 62.7%; Score 32; DB 1; Length 1247;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 CCGXXXC 10
    |||
Db 870 CCGTAEC 876

RESULT 12
POL_S_SFV STANDARD; PRT; 1253 AA.
AC P03315;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Structural polyprotein (PI30) [Contains: Coat protein C (EC 3.4.21.-)
DE (capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
DE 6 kDa peptide; Spike glycoprotein E1].
OS Semliki forest virus (SFV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=11033;
RN [1]
RP SEQUENCE OF 1-305 FROM N.A.
RX MEDLINE=81101055; PubMed=6935652;
RA Garoff H., Frischauf A.-M., Simons K., Lehrach H., Delius H.;
RT "The capsid protein of Semliki Forest virus has clusters of basic
RT amino acids and prolines in its amino-terminal region.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:6376-6380(1980).
RN [2]
RP SEQUENCE OF 266-1253 FROM N.A.
RX MEDLINE=81052444; PubMed=6985476;
RA Garoff H., Frischauf A.-M., Simons K., Lehrach H., Delius H.;
RT "Nucleotide sequence of cDNA coding for Semliki Forest virus membrane
RT glycoproteins.";
RL Nature 288:236-241(1980).
RN [3]
RP RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 119-267.
RX MEDLINE=97248723; PubMed=9094737;
RA Choi H.-K., Lu G., Lee S., Wengler G., Rossmann M.G.;
RT "Structure of Semliki Forest virus core protein.";
RL Proteins 27:345-359(1997).
CC -1- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THE 6 kDa POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
CC SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL
CC HEMAGGLUTININ.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X04129; CAA27742.1; AUT_SEQ.
PIR; A03915; VHWV
PDB; 1VCP; 07-DEC-96.
PDB; 1VCQ; 07-DEC-96.
MEROPS; S03.001; -
InterPro; IPR002548; Alpha_E1_glycop.
InterPro; IPR000936; Alpha_E2_glycop.
InterPro; IPR002533; Alpha_E3_glycop.
InterPro; IPR001836; Alpha_core.
InterPro; IPR000930; Togavirin.
Pfam; PF00944; Alpha_core; 1.
Pfam; PF01589; Alpha_E1_glycop; 1.
Pfam; PF00943; Alpha_E2_glycop; 1.
Pfam; PF01563; Alpha_E3_glycop; 1.
PRINTS; PR00798; TOGAVIRIN
Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;
Serine protease; 3D-structure.
CHAIN 1 267 COAT PROTEIN C.
FT CHAIN 268 333 SPIKE GLYCOPROTEIN E3.
FT CHAIN 334 755 SPIKE GLYCOPROTEIN E2.
FT CHAIN 756 815 6 KDA PEPTIDE.
FT CHAIN 816 1253 SPIKE GLYCOPROTEIN E1.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 151 151 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 219 219 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 119 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 533 533 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 956 956 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1253 AA; 138017 MW; 2A73228D08B82AC5 CRC64;

Query Match 62.7%; Score 32; DB 1; Length 1253;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGXXXC 10
DB 877 CCGASEC 883

RESULT 13
POL_S_RVNN STANDARD; PRT; 1254 AA.
AC P13890;
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Structural polyprotein (p130) [Contains: Coat protein C (EC 3.4.21.-)
DE (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
DE 6 kDa peptide; Spike glycoprotein E1].
OS Ross river virus (strain NB5092) (RRV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=11031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88179556; PubMed=2833022;
RA Faragher S.G., Meek A.D.J., Rice C.M., Dalgarno L.;
RT "Genome sequences of a mouse-avirulent and a mouse-virulent strain of
RT Ross River virus.";
RL Virology 163:509-526(1988).

CC -!- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: THE 6 kDa PEPTIDE PROBABLY SERVES AS THE SIGNAL
CC SEQUENCE FOR THE SPIKE GLYCOPROTEIN E1, WHICH IS THE VIRAL
CC HEMAGGLUTININ.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20162; AAA96330.1; -
CC PIR; B28605; VHWVRA.
CC HSSP; P03315; 1VCP.
CC MEROPS; S03.001; -
CC InterPro; IPR002548; Alpha_E1_glycop.
CC InterPro; IPR000936; Alpha_E2_glycop.
CC InterPro; IPR002533; Alpha_E3_glycop.
CC InterPro; IPR001836; Alpha_core.
CC InterPro; IPR000930; Togavirin.
CC Pfam; PF00944; Alpha_core; 1.
CC Pfam; PF01589; Alpha_E1_glycop; 1.
CC Pfam; PF00943; Alpha_E2_glycop; 1.
CC Pfam; PF01563; Alpha_E3_glycop; 1.
CC PRINTS; PR00798; TOGAVIRIN.
CC Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;
CC Serine protease.
CC CHAIN 1 270 COAT PROTEIN C.
CC CHAIN 271 334 SPIKE GLYCOPROTEIN E3.
CC CHAIN 335 756 SPIKE GLYCOPROTEIN E2.
CC CHAIN 757 816 6 KDA PEPTIDE.
CC CHAIN 817 1254 SPIKE GLYCOPROTEIN E1.
CC ACT_SITE 148 148 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 154 154 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 222 222 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CARBOHYD 281 281 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 957 957 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1254 AA; 138682 MW; 33D2D6DA589CE54B CRC64;

Query Match 62.7%; Score 32; DB 1; Length 1254;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 CCGXXXC 10
DB 878 CCGTSEC 884

RESULT 14
POL_S_RVNN STANDARD; PRT; 1254 AA.
AC P08431; Q88613; Q88614; Q88615; Q88616; Q88617;
DT 01-AUG-1988 (Rel. 08, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Structural polyprotein (p130) [Contains: Coat protein C (EC 3.4.21.-)
DE (Capsid protein C); Spike glycoprotein E3; Spike glycoprotein E2;
DE 6 kDa peptide; Spike glycoprotein E1].
OS Ross river virus (strain T48) (RRV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=11032;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83303839; PubMed=6310876;

RA Dalgarno L., Rice C.M., Strauss J.H.;
RT "Ross River virus 26 s RNA: complete nucleotide sequence and deduced
RL sequence of the encoded structural proteins.";
RN Virology 129:170-187(1983).
RP [2]
RX SEQUENCE OF 1-11 FROM N.A.
RA MEDLINE=8303346; PubMed=6291034;
RX Ou J.H., Rice C.M., Dalgarno L., Strauss E.G., Strauss J.H.;
RT "Sequence studies of several alphavirus genomic RNAs in the region
RL containing the start of the subgenomic RNA";
RN Proc. Natl. Acad. Sci. U.S.A. 79:5235-5239(1982).
RP [3]
RX SEQUENCE OF 335-756 FROM N.A.
RA MEDLINE=89073770; PubMed=2849242;
RX Burness A.T.H., Pardee I., Faragher S.G., Vratil S., Dalgarno L.;
RT "Genetic stability of Ross River virus during epidemic spread in
RL nonimmune humans";
RX Virology 167:639-643(1988).
CC -1- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
CC -1- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THE 6 kDa POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
CC SEQUENCE FOR THE SPIKE GLYCOPROTEIN E1, WHICH IS THE VIRAL
CC HEMAGGLUTININ.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K00046; AAA47404.1; -
DR EMBL; M23708; -; NOT_ANNOTATED_CDS.
DR RIR; A31833; VHWV48.
DR HSP; P03315; IVP.
DR MEMOPS; S03.001; -
DR InterPro; IPR002548; Alpha_E1_glycop.
DR InterPro; IPR000936; Alpha_E2_glycop.
DR InterPro; IPR002533; Alpha_E3_glycop.
DR InterPro; IPR000930; Togavirin.
DR Pfam; PF01589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
KW Coat protein; Polypeptide; Transmembrane; Glycoprotein; Hydrolase;
KW Serine protease.
FT CHAIN 1 270 COAT PROTEIN C.
FT CHAIN 271 334 SPIKE GLYCOPROTEIN E3.
FT CHAIN 335 756 SPIKE GLYCOPROTEIN E2.
FT CHAIN 757 816 6 KDA PEPTIDE.
FT CHAIN 817 1254 SPIKE GLYCOPROTEIN E1.
FT ACT_SITE 148 148 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 154 154 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 222 222 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 321 281 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 957 957 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1254 AA; 138467 MW; D3E5534919597E96 CRC64;

Query Match 62.7%; Score 32; DB 1; Length 1254;
Best Local Similarity 57.1%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 CCGXXXC 10
|||
Db 878 CCGTSEC 884

RESULT 15
TXM7_DENAN
ID TXM7_DENAN STANDARD; PRT; 65 AA.
AC P80970;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Muscarinic toxin 7 (MT-7).
OS Dendroaspis angusticeps [Eastern green mamba].
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Dendroaspis.
OX NCBI_TaxID=8618;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RA Jolkonen M.;
RL Thesis (1996), University of Uppsala, Sweden.
CC -1- FUNCTION: BINDS AND BLOCKS M1-TYPE MUSCARINIC ACETYLCHOLINE
CC RECEPTOR. IT IS THE MOST SELECTIVE M1-ACHR LIGAND KNOWN SO FAR.
CC -1- SUBUNIT: MONOMER.
DR HSP; P01382; INTN.
DR InterPro; IPR003571; Snake_toxin.
DR Pfam; PF00087; toxin; 1.
DR ProDom; PD000206; Snake_toxin; 1.
DR PROSITE; PS00272; SNAKE_TOXIN; 1.
KW Venom; Neurotoxin.
FT DISULFID 3 24 BY SIMILARITY.
FT DISULFID 17 42 BY SIMILARITY.
FT DISULFID 46 57 BY SIMILARITY.
FT DISULFID 58 63 BY SIMILARITY.
SQ SEQUENCE 65 AA; 7480 MW; 69D8FAAB7EC864B7 CRC64;

Query Match 60.8%; Score 31; DB 1; Length 65;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 CCGXXXC 10
|||
Db 57 CCGTDC 63

Search completed: May 19, 2002, 08:07:49
Job time: 12215 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 01:15:02 ; Search time 42.19 Seconds
(without alignments)
8.105 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCGXXXCXXCX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
-score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	78.4	460	1 US-08-689-974-5	Sequence 5, Appli
2	40	78.4	460	2 US-09-058-376-5	Sequence 5, Appli
3	40	78.4	1345	2 US-08-977-767-3	Sequence 3, Appli
4	39	76.5	57	1 US-07-609-716-56	Sequence 56, Appli
5	39	76.5	246	2 US-08-704-931-2	Sequence 2, Appli
6	38	74.5	38	1 US-07-609-716-53	Sequence 53, Appli
7	38	74.5	801	1 US-07-906-349A-6	Sequence 6, Appli
8	37	72.5	41	5 PCT-US96-01720-7	Sequence 7, Appli
9	37	72.5	55	3 US-08-476-509B-46	Sequence 46, Appli
10	37	72.5	1417	4 US-08-900-230-3	Sequence 3, Appli
11	36	70.6	21	2 US-08-754-431A-10	Sequence 37, Appli
12	35	68.6	1400	4 US-08-630-915A-37	Sequence 10, Appli
13	33	64.7	22	1 US-08-423-399B-27	Sequence 27, Appli
14	33	64.7	50	4 US-08-900-230-8	Sequence 8, Appli
15	33	64.7	143	4 US-08-990-823-112	Sequence 112, App
16	32	62.7	27	4 US-09-073-407-1	Sequence 1, Appli
17	32	62.7	27	4 US-09-073-407-2	Sequence 2, Appli
18	32	62.7	27	4 US-09-073-407-3	Sequence 3, Appli
19	32	62.7	27	4 US-09-073-407-4	Sequence 4, Appli
20	32	62.7	27	4 US-09-073-407-5	Sequence 5, Appli
21	32	62.7	27	4 US-09-073-407-6	Sequence 6, Appli
22	32	62.7	27	4 US-09-073-407-7	Sequence 7, Appli
23	32	62.7	27	4 US-09-073-407-8	Sequence 8, Appli
24	32	62.7	43	6 5212296-3	Patent No. 5212296
25	32	62.7	47	1 US-08-451-947-94	Sequence 94, Appli
26	32	62.7	47	2 US-08-424-826A-94	Sequence 94, Appli
27	32	62.7	47	3 US-08-482-085B-91	Sequence 91, Appli

28	32	62.7	47	3 US-08-928-694-94	Sequence 94, Appli
29	32	62.7	47	5 PCT-US91-06950-94	Sequence 94, Appli
30	32	62.7	64	6 5212296-10	Patent No. 5212296
31	32	62.7	109	2 US-08-527-044-2	Sequence 2, Appli
32	32	62.7	109	3 US-09-013-780-2	Sequence 2, Appli
33	32	62.7	111	2 US-08-485-937-4	Sequence 4, Appli
34	32	62.7	111	2 US-08-373-215-4	Sequence 4, Appli
35	32	62.7	111	5 PCT-US93-06552-4	Sequence 4, Appli
36	32	62.7	141	2 US-08-485-937-1	Sequence 1, Appli
37	32	62.7	141	2 US-08-373-215-1	Sequence 1, Appli
38	32	62.7	141	5 PCT-US93-06552-1	Sequence 1, Appli
39	32	62.7	1253	1 US-07-920-281C-3	Sequence 3, Appli
40	32	62.7	1253	4 US-08-466-277-3	Sequence 3, Appli
41	31.5	61.8	51	3 US-08-482-085B-90	Sequence 30, Appli
42	31	60.8	23	4 US-08-900-230-30	Sequence 30, Appli
43	31	60.8	24	1 US-08-175-155-43	Sequence 43, Appli
44	31	60.8	24	2 US-08-643-034A-6	Sequence 6, Appli
45	31	60.8	24	3 US-08-482-085B-78	Sequence 78, Appli

ALIGNMENTS

RESULT 1
US-08-689-974-5
; Sequence 5, Application US/08689974
; Patent No. 5776732
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Murray, Lynn E.
; TITLE OF INVENTION: NOVEL HUMAN INDUCED TUMOR PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/689,974
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0113 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 310100
US-08-689-974-5

Query Match 78.4%; Score 40; DB 1; Length 460;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

Db 13 CCGSAGCALC 22

RESULT 2

US-09-058-376-5
; Sequence 5, Application US/09058376
; Patent No. 6080841
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Murray, Lynn E.
; TITLE OF INVENTION: NOVEL HUMAN INDUCED TUMOR PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,376
; FILING DATE:
; PRIOR APPLICATION DATA:
; -APPLICATION NUMBER: US/08/689,974
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0113 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 310100
US-09-058-376-5

Query Match 78.4%; Score 40; DB 3; Length 460;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXCXC 13
111 1 1

Db 13 CCGSAGCALC 22

RESULT 3

US-08-977-767-3
; Sequence 3, Application US/08977767
; Patent No. 5972684
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Greenwald, Sara
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: CARBONIC ANHYDRASE VIII
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,767
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0423 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1345 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1532042
US-08-977-767-3

Query Match 78.4%; Score 40; DB 2; Length 1345;
Best Local Similarity 50.0%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXXCXC 13
111 1 1

Db 1089 CCGTACGCGC 1098

RESULT 4

US-07-609-716-56
; Sequence 56, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: A-55186-3/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-07-609-716-56

Query Match 76.5%; Score 39; DB 1; Length 57;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXXC 13
||| |
DB 28 CCGGGCTAC 37

RESULT 5
US-08-704-931-2
Sequence 2, Application US/08704931
Patent No. 5885797
GENERAL INFORMATION:
APPLICANT: Chen, Chao-Min (Amy)
APPLICANT: Kraut, No. 5885797/bert
APPLICANT: Groudiue, Mark
APPLICANT: Weintraub, Harold
TITLE OF INVENTION: No. 5885797el DNA Sequences Encoding Proteins
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Stratton Ballew, PLLC
STREET: 1218 Third Avenue, Suite 1313
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/704,931
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: King, Jeffrey J
REGISTRATION NUMBER: 38,515
REFERENCE/DOCKET NUMBER: HUI1.P02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-683-1496
TELEFAX: 206-682-0446
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 246 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-704-931-2

Query Match 76.5%; Score 39; DB 2; Length 246;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXXC 13
||| |
DB 203 CCGSGECADC 212

RESULT 6
US-07-609-716-53
Sequence 53, Application US/07609716
Patent No. 5514581
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Cappello, Joseph
TITLE OF INVENTION: Functional Recombinantly Prepared
TITLE OF INVENTION: Synthetic Protein Polymer
NUMBER OF SEQUENCES: 118
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,716
FILING DATE: 06-NOV-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: A-55186-3/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-07-609-716-53

Query Match 74.5%; Score 38; DB 1; Length 38;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXXC 13
||| |
DB 1 CCGGGTCTC 10

RESULT 7
US-07-906-349A-6
Sequence 6, Application US/07906349A
Patent No. 5434064
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Skolnik, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES
TITLE OF INVENTION: TARGET PROTEINS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W.

;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20004
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/906,349A
;; FILING DATE: 30-JUN-1992
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/643,237
;; FILING DATE: 18-JAN-1991
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 801 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-07-906-349A-6

Query Match 74.5%; Score 38; DB 1; Length 801;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
||| |
DB 341 CCGGTCAC 350

RESULT 8
PCT-US96-01720-7
;; Sequence 7, Application PC/TUS9601720
;; GENERAL INFORMATION:
;; APPLICANT:
;; TITLE OF INVENTION: MODIFIED-AFFINITY STREPTAVIDIN
;; NUMBER OF SEQUENCES: 11
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US96/01720
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/387,055
;; FILING DATE: 09-FEB-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Parmelee, Steven W.
;; REGISTRATION NUMBER: 31,990
;; REFERENCE/DOCKET NUMBER: 16336-5PC
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 41 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; PCT-US96-01720-7

Query Match 72.5%; Score 37; DB 5; Length 41;
Best Local Similarity 50.0%; Pred. No. 18;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 CCGXXCXC 13
||| |
DB 16 CCGGTCAC 25

RESULT 9
US-08-476-509B-46
;; Sequence 46, Application US/08476509B
;; Patent No. 6034212
;; GENERAL INFORMATION:
;; APPLICANT: SUDOL, MARIUS
;; APPLICANT: PEER, BORK
;; APPLICANT: HENRY, CHEN
;; TITLE OF INVENTION: A SH2 DOMAIN ASSOCIATED PROTEIN, A
;; SIGNALING DOMAIN THEREIN, NUCLEIC ACIDS ENCODING THE
;; TITLE OF INVENTION: PROTEIN AND THE DOMAIN, AND DIAGNOSTIC AND THERAPEUTICS
;; TITLE OF INVENTION: THEREOF
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Klauber & Jackson
;; STREET: 411 Hackensack Avenue
;; CITY: Hackensack
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07601

COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/476,509B
;; FILING DATE: 01-DEC-1994
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Jackson Esq., David A.
;; REGISTRATION NUMBER: 26,742
;; REFERENCE/DOCKET NUMBER: 600-1-101 CIP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201 487-5800
;; TELEFAX: 201 343-1684
;; TELEX: 133521

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:
;; LENGTH: 55 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "PRIMER"
;; HYPOTHETICAL: NO
;; US-08-476-509B-46

Query Match 72.5%; Score 37; DB 3; Length 55;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
||| |
DB 11 CCGGTCAC 20

RESULT 10
US-08-900-230-3
;; Sequence 3, Application US/08900230
;; Patent No. 6329197
;; GENERAL INFORMATION:
;; APPLICANT: Bard, Jonathan A.
;; TITLE OF INVENTION: DNA ENCODING GALANIN GALR3 RECEPTORS AND
;; TITLE OF INVENTION: USES THEREOF

; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham LLP
 ; STREET: 1185 Avenue of The Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 11036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/900,230
 ; FILING DATE: 23-JUL-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-278-0400
 ; TELEFAX: 212-391-0525
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1417 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: NO
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; US-08-900-230-3

 Query Match 72.5%; Score 37; DB 4; Length 1417;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

 QY 4 CCGXXXCXXC 13
 ||| |
 Db 960 CCGCTTCGCC 969

 RESULT 11
 US-08-754-431A-10
 ; Sequence 10, Application US/08754431A
 ; Patent No. 5824512
 ; GENERAL INFORMATION:
 ; APPLICANT: Pazirandeh, Mehran
 ; APPLICANT: Campbell, James R.
 ; TITLE OF INVENTION: BACTERIA EXPRESSING METALLOTHIONEIN GENE INTO
 ; TITLE OF INVENTION: THE PERIPLASMIC SPACE AND METHOD OF USING SUCH
 ; TITLE OF INVENTION: BACTERIA IN ENVIRONMENTAL CLEANUP
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Naval Research Laboratory
 ; STREET: 4555 Overlook Avenue, S.W.
 ; CITY: Washington, D.C.
 ; STATE: not applicable
 ; COUNTRY: United States of America
 ; ZIP: 20375-5000
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette-3.5 inch, 1.44 Mb storage
 ; COMPUTER: IBM AT compatible
 ; OPERATING SYSTEM: WINDOWS 95
 ; SOFTWARE: Word Perfect 8
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/754,431A
 ; FILING DATE: 22-NOV-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA: not applicable

; ATTORNEY/AGENT INFORMATION:
 ; NAME: EDELBERG, BARRY ALLEN
 ; REGISTRATION NUMBER: 31,012
 ; REFERENCE/DOCKET NUMBER: NC 77,640
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)404-1558
 ; TELEFAX: (202)404-7380
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 21 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-754-431A-10

 Query Match 70.6%; Score 36; DB 2; Length 21;
 Best Local Similarity 50.0%; Pred. No. 17;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

 QY 4 CCGXXXCXXC 13
 ||| |
 Db 3 CCGKGHGCC 12

 RESULT 12
 US-08-630-915A-37
 ; Sequence 37, Application US/08630915A
 ; Patent No. 6309820
 ; GENERAL INFORMATION:
 ; APPLICANT: SPARKS, Andrew B.
 ; APPLICANT: HOFFMAN, No. 6309820h
 ; APPLICANT: KAY, Brian K.
 ; APPLICANT: FOWLKES, Dana M.
 ; APPLICANT: MCCONNELL, Stephen J.
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
 ; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
 ; NUMBER OF SEQUENCES: 227
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/630,915A
 ; FILING DATE: 03-APR-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 1101-174
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 37:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1400 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; US-08-630-915A-37

 Query Match 68.6%; Score 35; DB 4; Length 1400;

Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
|| | |
Db 153 CCATTCTGC 162

RESULT 13
US-08-423-399B-27
; Sequence 27, Application US/08423399B
; Patent No. 5663314
; GENERAL INFORMATION:
; APPLICANT: Seger, R., Seger, D., Ahn, N.G., and Krebs, E.G.
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION MAPK KINASE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-5.25 inch, 1.2Mb storage
; COMPUTER: IBM PC/386 Compatible
; OPERATING SYSTEM: MS-DOS 4.01
; SOFTWARE: Word for Windows-t
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/423,399B
; FILING DATE: April 18, 1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/980,608
; FILING DATE: No. 5663314ember 20, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sheiness, Diana K.
; REGISTRATION NUMBER: 35,356
; REFERENCE/DOCKET NUMBER: WRFO-1-8427
; TELEPHONE: 1-206-682-8100; 1-206-224-0712 (direct)
; TELEFAX: 1-206-224-0779
; TELEX: 4938023
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; DESCRIPTION: 56A sequence for construction of
; DESCRIPTION: degenerate PCR primers; Table 4
US-08-423-399B-27

Query Match 64.7%; Score 33; DB 1; Length 22;
Best Local Similarity 40.0%; Pred. No. 44;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
|| | |
Db 7 CCNRYTCYC 16

RESULT 14
US-08-900-230-8
; Sequence 8, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANIN GPCR RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of The Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 11036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,230
FILING DATE: 23-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-900-230-8

Query Match 64.7%; Score 33; DB 4; Length 50;
Best Local Similarity 40.0%; Pred. No. 70;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CCGXXCXC 13
|| | |
Db 14 CCTGTCATC 23

RESULT 15
US-08-990-823-112
; Sequence 112, Application US/08990823D
; Patent No. 6228371
; GENERAL INFORMATION:
; APPLICANT: Nano, Francis
; TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding
; TITLE OF INVENTION: Immunostimulatory Peptides
; FILE REFERENCE: 49086
; CURRENT APPLICATION NUMBER: US/08/990,823D
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: US 96/10375
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,254
; EARLIER FILING DATE: 1995-06-15
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 112
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-08-990-823-112

Query Match 64.7%; Score 33; DB 4; Length 143;
Best Local Similarity 44.4%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CGXXCXC 13

Db 63 CGAATCTTC 71

Search completed: May 19, 2002, 04:44:12
Job time: 12550 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 01:14:13 ; Search time 86.36 seconds
(without alignments)
18,006 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCGGXXCX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues 747574
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SID55/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	78.4	210	AAW85717	Novel protein (Clo
2	40	78.4	1679	AAU07343	1-aminocyclopropan
3	39	76.5	31	AA123346	Peptide sequence u
4	39	76.5	31	AA119099	4-Comarate CoA li
5	39	76.5	31	AAE05833	Liquidambar sp. an
6	39	76.5	169	AAAB43505	Human cancer assoc
7	39	76.5	246	AAW53007	Mus musculus I-mfa
8	39	76.5	919	AAAG84970	Shrimp white spot
9	39	76.5	1231	ABBS9738	Drosophila melanog
10	38	74.5	80	AAU43831	Propionibacterium
11	38	74.5	153	AA170234	Human RNA-associat

12	38	74.5	237	22	ABB69415	Drosophila melanog
13	38	74.5	973	22	AAW41045	Human polypeptide
14	38	74.5	1023	22	AAW11436	D. discoideum ster
15	38	74.5	1278	22	AAW39259	Human polypeptide
16	37	72.5	166	22	AAU41699	Propionibacterium
17	37	72.5	225	21	AAW36084	Neospora caninum N
18	37	72.5	250	22	AAW79002	Human protein SEQ
19	37	72.5	301	22	AAW79986	Human protein SEQ
20	37	72.5	369	22	ABG15487	Novel human diagno
21	36	70.6	12	21	AAW08014	Generic formula fo
22	36	70.6	12	21	AAW08015	Generic formula fo
23	36	70.6	12	21	AAW08018	Amino acid sequenc
24	36	70.6	12	21	AAW08019	Amino acid sequenc
25	36	70.6	13	21	AAW08012	Generic formula fo
26	36	70.6	13	21	AAW08013	Generic formula fo
27	36	70.6	13	21	AAW08017	Amino acid sequenc
28	36	70.6	13	21	AAW92229	Chi-conotoxin pept
29	36	70.6	13	21	AAW92230	Chi-conotoxin pept
30	36	70.6	21	19	AAW69478	Addition sequenc
31	36	70.6	61	21	AAW08016	Amino acid sequenc
32	36	70.6	61	21	AAW92231	Chi-conotoxin, chi
33	35	68.6	55	22	ABW66559	Drosophila melanog
34	34	66.7	149	21	AAW34801	Arabidopsis thalia
35	34	66.7	463	22	ABG28169	Novel human diagno
36	33	64.7	51	20	AAW30722	Amino acid sequenc
37	33	64.7	52	21	AAW45580	Arabidopsis thalia
38	33	64.7	130	22	AAW85645	Human immune/haema
39	33	64.7	412	21	AAW43596	Arabidopsis thalia
40	32	62.7	25	22	AAW64882	Murine peptide SEQ
41	32	62.7	60	22	AAW64881	Murine protein SEQ
42	32	62.7	64	12	AAW11804	Iron sulphur prote
43	32	62.7	84	22	AAW64880	Murine protein SEQ
44	32	62.7	100	22	AAW87161	Human immune/haema
45	32	62.7	109	17	AAW84086	T-lymphocyte stimu

ALIGNMENTS

RESULT 1
AAW85717
ID AAW85717 standard; Protein; 210 AA.
XX
XX AAW85717:
AC
XX 27-SEP-1999 (first entry)
DT
XX
XX Novel protein (Clone AC41_1).
DE

XX Polynucleotide; protein; nutrition; cytokine; cell proliferation;
KW cell differentiation; immunostimulation; immunosuppression;
KW haematopoiesis regulation; tissue growth; activin; inhibin;
KW chemotaxis; chemokinesis; haemostasis; thrombolysis; receptor;
KW ligand; anti-inflammatory; tumour suppression; gene therapy.

XX Homo sapiens.
XX
XX WO9920644-Al.
XX
XX 29-APR-1999.
PD
XX
XX 16-OCT-1998; 98WO-US22034.
PF
XX
XX 18-OCT-1997; 97US-0955557.
XX
XX (GEM) GENETICS INST INC.
PA
XX
XX Agostino MJ, Bowman MR, Evans C, Jacobs K, Lavallie ER;
PI McCoy JM, Merberg D, Racie LA, Spaulding V, Treacy M;
XX
XX WPI; 1999-288272/24.
XX
XX N-PSDB; AAX08682.

PT New polynucleotides encoding secreted human proteins

XX Claim 8; Page 100; 136pp; English.

CC The new human secreted proteins are encoded by polynucleotides
 CC obtained from human placenta, adult testes, fetal kidney, fetal
 CC brain, adult brain, adult brain and adult blood cDNA libraries.
 CC The polynucleotides and proteins are predicted to have biological
 CC activities which would make them suitable for treating, preventing or
 CC ameliorating medical conditions in humans and animals. Suggested
 CC activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, and tumour inhibition
 CC activity. The polynucleotides are also stated to be useful for gene
 CC therapy. The sequences identified by a secretory leader
 CC sequence motif in the polynucleotide and it is thought that the
 CC encoded proteins have biological activity by virtue of their secreted
 CC nature. This polypeptide was encoded by a clone designated AC41_1.
 CC (See AAX08682).

XX Sequence 210 AA;

Query Match 78.4%; Score 40; DB 20; Length 210;

Best Local Similarity 50.0%; Pred. No. 1.2e+02; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

Db 13 ccgsagcslc 22

RESULT 2

AAU07343
 ID AAU07343 standard; Protein; 1679 AA.

XX AC AAU07343;

XX DT 04-DEC-2001 (first entry)

XX DE 1-aminocyclopropane carboxylate (ACPC) synthase #12.

XX KW 1-aminocyclopropane carboxylate synthase; ACPC synthase; brain trauma;
 KW excitatory neurotransmission; neurodegeneration; stroke; nerve damage;
 KW neurodegenerative disease; Alzheimer's disease; depression; epilepsy;
 KW alcohol abuse; cognitive function; memory; learning impairment; human.

XX OS Homo sapiens.

XX PN WO200168879-A2.

XX PD 20-SEP-2001.

XX PF 14-MAR-2001; 2001WO-EP02857.

XX PR 14-MAR-2000; 2000US-0189086.

XX PR 03-APR-2000; 2000US-0194702.

XX PA (FARB) BAYER AG.

XX PI Ramakrishnan S;

XX XX WPI; 2001-550286/61.

XX Isolated polynucleotide encoding a human

PT 1-aminocyclopropane-carboxylate (ACPC) synthase, useful for treating
 PT brain trauma and neurodegenerative disease (e.g. Alzheimer's disease,
 PT depression, epilepsy) -

XX

PS Claim 1; Page 231-237; 242pp; English.

XX The invention relates to reagents and methods for regulating excitatory
 CC neurotransmission, and to prevent neurodegeneration. The method involves
 CC the use of an expression vector or a reagent that modulates the activity
 CC of a 1-aminocyclopropane-carboxylate (ACPC) synthase polypeptide. The
 CC reagent is useful for modulating the activity of an ACPC synthase in a
 CC disease such as stroke, a nerve damage or a neurodegenerative disease.
 CC The ACPC synthase polypeptide, polynucleotides and modulators are also
 CC useful for treating brain trauma and neurodegenerative disease (e.g.
 CC Alzheimer's disease, depression, epilepsy). The ACPC synthase modulators
 CC are also useful for treating alcohol abuse and improve cognitive function
 CC and memory of patients with learning impairment. The present sequence
 CC represents the amino acid sequence of human 1-aminocyclopropane-
 CC carboxylate (ACPC) synthase #12, used in the method of the invention.

XX Sequence 1679 AA;

Query Match 78.4%; Score 40; DB 22; Length 1679;

Best Local Similarity 50.0%; Pred. No. 3.8e+02; Mismatches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

Db 1112 ccgtggcttc 1121

RESULT 3

AA23346
 ID AAY23346 standard; Peptide; 31 AA.

XX AC AAY23346;

XX DT 01-SEP-1999 (first entry)

XX DE Peptide sequence used to derive a primer for CL gene isolation.

XX KW Coumarate CoA ligase; syringyl lignin; angiosperm; gymnosperm;
 KW loblolly pine; Pinus taeda; lignin; woody plant; paper manufacture;
 KW delignification; pulp mill.

XX OS Unidentified.

XX PN WO9931243-A1.

XX PD 24-JUN-1999.

XX PF 16-DEC-1998; 98WO-US26784.

XX PR 16-DEC-1997; 97US-0991677.

XX PA (INTO) INT PAPER CO.

XX PI Carraway DT, Chiang VL, Smeltzer RH;

XX XX WPI; 1999-405034/34.

XX New DNA sequences encoding angiosperm enzymes involved in
 PT biosynthesis of syringyl lignin

XX Example 1; Page 11; 83pp; English.

XX AAY23346-48 represent peptide sequences used to derived degenerate
 CC primers for coumarate CoA ligase (CL) gene isolation. CL is
 CC involved in syringyl lignin production in an angiosperm. The
 CC sequences are used to produce syringyl lignin in gymnosperms,
 CC especially the loblolly pine (Pinus taeda). It is necessary to
 CC remove much of the lignin from the fiber/lignin network of woody
 CC plants in paper manufacture. Greater proportions of syringyl lignin
 CC result in a higher delignification rate and hence a more efficient
 CC pulp mill operation.

XX

SQ Sequence 31 AA;

Query Match 76.5%; Score 39; DB 20; Length 31;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 7 ccgggiaciac 16

RESULT 4
AAB19699
ID AAB19699 standard; Peptide; 31 AA.

AC AAB19699;
XX
DT 05-FEB-2001 (first entry)

DE 4-Coumarate CoA ligase conserved peptide.

XX Sweetgum; angiosperm; 4-coumarate CoA ligase; conifer;
KW loblolly pine; transgenic plant; lignin; paper; pulping.
XX Planta.

OS
XX WO200058489-A2.

PN
XX 05-OCT-2000.

PD
XX 24-MAR-2000; 2000WO-0508083.

PF
XX 25-MAR-1999; 99US-0277248.

PR (INTO) INT PAPER CO.

PA Chiang VL, Carraway DT;

XX WPI; 2000-647240/62.

XX Use of angiosperm coniferyl aldehyde 5-hydroxylase which catalyzes
PT 5-hydroxylation of coniferyl aldehyde, for modifying lignin
PT biosynthesis in gymnosperms, involves expressing the enzyme in a
PT gymnosperm plant -

XX Example 1; Page 19; 123pp; English.

XX The present sequence is that of a conserved peptide of known
CC 4-coumarate Co ligases (4CL). A primer based on this sequence
CC was used in the identification of cDNA (see AAB8684) encoding
CC Sweetgum 4CL (see AAB19699), an enzyme involved in lignin
CC biosynthetic pathways. An aim of the invention is to identify,
CC sequence and clone specific genes from an angiosperm such as
CC Sweetgum that are involved in production of syringyl lignin, and
CC to then introduce such genes into the genome of a gymnosperm, such
CC as loblolly pine, to induce production of syringyl lignin and
CC thereby provide enhanced pulpability to the wood structure.

XX Sequence 31 AA;

Query Match 76.5%; Score 39; DB 21; Length 31;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 7 ccgggiaciac 16

RESULT 5
AAB05833

ID AAE05833 standard; peptide; 31 AA.

XX AAE05833;

AC 24-SEP-2001 (first entry)

XX Liquidambar sp. angiosperm 4-coumarate CoA ligase conserved peptide #1.

DE Sweetgum; angiosperm; 4-coumarate CoA ligase; 4CL;
KW syringyl lignin; gymnosperm; pulp; papermaking; guaiacyl lignin;
KW delignification.

XX Liquidambar sp.

XX US6252135-B1.

PN 26-JUN-2001.

XX 16-DEC-1997; 97US-0991677.

XX 16-DEC-1996; 96US-0033381.

XX (INTO) INT PAPER CO.

XX Chiang VL, Carraway DT, Smeltzer RH;

XX WPI; 2001-456503/49.

XX New P450-2 DNA encoding an enzyme involved in the biosynthesis of
PT syringyl lignin monomer units, for inducing the production of syringyl
PT lignin in gymnosperms and improving delignification for the production
PT of pulp -

XX Example 1; Column 45-46; 59pp; English.

XX The invention relates to angiosperm sweetgum proteins bifunctional-O-
CC methyl transferase (bi-OMT), 4-coumarate CoA ligase (4CL), ferulic acid
CC 5-hydroxylase (P450-1 and P450-2) and their corresponding DNA molecules.
CC Angiosperm sweetgum proteins are useful for inducing the production of
CC syringyl lignin in gymnosperms for improved delignification in the
CC production of pulp for papermaking and other applications. The invention
CC also provides a method for modifying genes involved in lignin
CC biosynthesis in loblolly pine gymnosperms which involves cloning the
CC sweetgum DNA and fusing it to a constitutive promoter to form an
CC expression cassette. The expression cassette is then inserted into the
CC gymnosperm genome, so that the production of syringyl lignin is increased
CC while the production of guaiacyl lignin (less preferred form) is
CC suppressed. The present sequence is Liquidambar sp. angiosperm
CC 4-coumarate CoA ligase (4CL) conserved peptide which is used for
CC designing RFLS sense PCR primer.

CC Note: This sequence is stated as being the same as that shown as
CC SEQ ID NO: 14 in column 9 of the specification where the amino acids
CC has been designated as 'Lie' for 'ile', 'Thlr' for 'Thr' and
CC 'Mrg' for 'Arg'.

XX Sequence 31 AA;

Query Match 76.5%; Score 39; DB 22; Length 31;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| |
Db 7 ccgggiaciac 16

RESULT 6
AAB43505
ID AAB43505 standard; Protein; 169 AA.

XX AAB43505;

XX

DT 08-FEB-2001 (first entry)
 XX Human cancer associated protein sequence SEQ ID NO:950.
 DE
 XX
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
 KW antidiabetic; antithymic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antihistoid; antiallergic; antithrombotic; cardiant;
 KW dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 XX Homo sapiens.
 XX
 XX WO200055350-A1.
 PN
 XX
 XX 21-SEP-2000.
 PD
 XX
 XX 08-MAR-2000; 2000WO-US05882.
 PF
 XX
 XX 12-MAR-1999; 99US-0124270.
 PR
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX Rosen CA, Ruben SM;
 PI
 XX
 XX WPI: 2000-587533/55.
 DR N-PSDB; AAC77714.
 DR
 XX
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 PT
 XX
 XX Claim 11; Page 1516; 2352pp; English.
 PS
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerary; immunomodulator;
 CC antidiabetic; antithymic; antirheumatic; antiarthritic;
 CC antiinflammatory; antihistoid; antiallergic; antithrombotic; antiviral;
 CC dermatologic; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 XX Sequence 169 AA;
 SQ
 Query Match 76.5%; Score 39; DB 21; Length 169;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 4 CCGXXXCXC 13
 ||| | |
 Db 126 ccgsgccadc 135
 RESULT 7
 AAW53007

ID AAW53007 standard; Protein; 246 AA.
 XX
 AC AAW53007;
 AC
 XX
 DT 03-AUG-1998 (first entry)
 DT
 XX Mus musculus I-mfa protein.
 DE
 XX I-mfa; inhibitor of MyoD family; treatment; diagnosis; myogenesis;
 KW defects; abnormal development; disease; cleidocranial dysplasia;
 KW CDD; rhabdomyosarcoma; muscle tissue cancer.
 KW
 XX Mus musculus.
 OS
 XX
 XX WO9808860-A1.
 PN
 XX
 XX 05-MAR-1998.
 PD
 XX
 XX 21-AUG-1997; 97WO-US14780.
 PF
 XX
 XX 27-AUG-1996; 96US-0704931.
 PR
 XX
 XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.
 PA
 XX
 XX Chen CMA, Groudine M, Kraut N, Weintraub H;
 PI
 XX
 XX WPI: 1998-179377/16.
 DR N-PSDB; AAV21282.
 DR
 XX
 XX Inhibitor of MyoD family proteins - useful for, e.g. treatment and
 PT diagnosis of defects in myogenesis responsible for abnormal
 PT development
 PT
 XX
 PS Disclosure; Pages 74-75; 92pp; English.
 PS
 XX
 CC The sequence is that of murine I-mfa (inhibitor of MyoD family)
 CC protein. Probes from the gene sequence can be used for determining the
 CC presence of an I-mf protein or analogue, or for detecting I-mf
 CC agonist activity in a test substance. The sequence can be used
 CC to treat, model and diagnose defects in myogenesis responsible
 CC for abnormal development and disease conditions in mammals,
 CC including humans. These include the human disease cleidocranial
 CC dysplasia (CDD), rhabdomyosarcomas and other cancers affecting
 CC muscle tissue in mammals.
 XX
 SQ Sequence 246 AA;
 Query Match 76.5%; Score 39; DB 19; Length 246;
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 4 CCGXXXCXC 13
 ||| | |
 Db 203 ccgsgccadc 212
 RESULT 8
 AAG84970
 ID AAG84970 standard; Protein; 919 AA.
 XX
 AC AAG84970;
 AC
 XX
 DT 11-SEP-2001 (first entry)
 DT
 XX Shrimp white spot Bacilliform virus (WSBV) protein 61.
 DE
 XX Shrimp white spot Bacilliform virus; WSBV; diagnosis; viral infection;
 KW antiviral agent; gene expression; antisense construct;
 KW transgenic viral resistant shrimp.
 KW
 XX White spot syndrome virus.
 OS
 XX

PN WO200138351-A2.
XX 31-MAY-2001.
PD
XX
XX 08-NOV-2000; 2000WO-US28888.
PF
XX 24-NOV-1999; 99CN-0124717.
XX
XX (PENY-) PE CORP NY.
PA (THUR-) THIRD INST OCEANOGRAPHY STATE OCEANI C A.
PA (SINO-) SINOGENOMAX CO LTD.
XX
PI Xu X, Yang F, He J, Pham L, He M, Ye Y, Shen Y, Kodira C;
XX
XX WPI; 2001-355877/37.
DR N-PSDB; AAH62750.
XX
XX Primary nucleotide sequence of the shrimp white spot Bacilliform virus
PT (WSBV), useful for producing viral polypeptides that can be used to
PT screen for agents that are useful for treating WSBV infection -
XX
XX Claim 1; Figure 3; 626pp; English.
PS
XX The invention provides the primary nucleotide sequence of the WSBV genome
CC (AAH62689), predicted transcript sequences (AAH62689-AAH62839) and
CC encoded proteins (AAG84910-AAG85051) and oligonucleotide sequences
CC (AAH62840-63160) suitable for use as primers or probes. The nucleic acid
CC molecules and proteins of the invention are useful for diagnosis and
CC monitoring viral infection, in screens for antiviral agents and for
CC monitoring viral gene expression or activity during a treatment regimen.
CC The nucleic acid molecules are also useful as antisense constructs to
CC control viral gene expression in infected cells and tissues and to create
CC transgenic viral resistant shrimp.
XX
XX Sequence 919 AA;
SQ

Query Match 76.5%; Score 39; DB 22; Length 919;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
Db 412 ccgagkicsec 421
||| | |

RESULT 9
ABB59738
ID ABB59738 standard; Protein; 1231 AA.
XX
XX ABB59738;
AC
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 6006.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
PR
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI

XX WPI; 2001-656860/75.
DR N-PSDB; ABL03841.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 6006; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1231 AA;
SQ

Query Match 76.5%; Score 39; DB 22; Length 1231;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
Db 232 ccgssfcddc 241
||| | |

RESULT 10
AAU43831
ID AAU43831 standard; Protein; 80 AA.
XX
XX AC AAU43831;
AC
XX 13-FEB-2002 (first entry)
DT
XX Propionibacterium acnes immunogenic protein #4727.
DE
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX WO200181581-A2.
PN
XX 01-NOV-2001.
PD
XX 20-APR-2001; 2001WO-US12865.
PF
XX 21-APR-2000; 2000US-199047P.
PR
XX 02-JUN-2000; 2000US-208841P.
PR
XX 07-JUL-2000; 2000US-216747P.
PR
XX (CORI-) CORIXA CORP.
PA
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59521.
DR
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX

PS Example 1; SEQ ID NO 5026; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 80 AA;

Query Match 74.5%; Score 38; DB 22; Length 80;

Best Local Similarity 50.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

Db 17 ccggaccsgc 26

RESULT 11

AAU70234

ID AAY70234 standard; Protein; 153 AA.

XX AC AAY70234;

XX 06-JUN-2000 (first entry)

DE Human RNA-associated protein-15 (RNAAP-15).

XX RNA-associated protein; RNAAP; human; clone 1879135; cytostatic;
KW immunosuppressive; antiinflammatory; keratolytic; neuroprotective;
KW antiarteriosclerotic; hepatotropic; antipsoriatic; virucide; anti-HIV;
KW antiallergic; antirheumatic; antiarthritic; ophthalmological; autoimmune;
KW antimicrobial; cell proliferative disorder; inflammation; cirrhosis;
KW actinic keratosis; bursitis; arteriosclerosis; artherosclerosis;
KW hepatitis; myelofibrosis; primary thrombocythemia; psoriasis; cancer;
KW mixed connective tissue disease; MCTD; HIV; uveitis; Crohn's disease;
KW allergy; rheumatoid arthritis; parasitic infection.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 5 /note= "Potential phosphorylation site"

FT Domain 26..43 /label= Zinc_finger

FT Modified-site 43 /note= "Potential phosphorylation site"

FT Modified-site 59 /note= "Potential phosphorylation site"

FT Modified-site 82 /note= "Potential phosphorylation site"

FT Modified-site 112 /note= "Potential phosphorylation site"

FT Modified-site 113 /note= "Potential phosphorylation site"

FT Modified-site 114 /note= "Potential phosphorylation site"

FT Modified-site /note= "Potential phosphorylation site"

FT 145 /note= "Potential phosphorylation site"

XX WO200011171-A2.

XX 02-MAR-2000.

XX 20-AUG-1999; 99WO-US19361.

XX 21-AUG-1998; 98US-0097550.

XX 12-JAN-1999; 99US-0115639.

XX (INCY-) INCYTE PHARM INC.

XX Hillman JL, Yue H, Tang YT, Corley NC, Guegler KJ, Gorgone GA;

XX Patterson C, Baughn MR, Lai P, Bandman O, Reddy R, Azimzai Y;

XX Shih LL, Yang J, Lu DAM;

XX WPI; 2000-237651/20.

XX N-PSDB; AAZ51264.

XX Human RNA-associated proteins useful in diagnosing, treating and
PT preventing cell proliferative, autoimmune, inflammatory and infectious
PT disorders

XX Claim 1; Page 94-95; 123pp; English.

XX The present amino acid sequence is the human RNA-associated protein-15
CC (RNAAP-15), identified in Incyte clone 1879135, derived from LEUKNOT03
CC library. It is expressed in reproductive, dermatologic, gastrointestinal,
CC musculoskeletal and haematopoietic/immune tissues. It has cytostatic,
CC immunosuppressive, antiinflammatory, antiarteriosclerotic, hepatotropic,
CC keratolytic, neuroprotective, antipsoriatic, anti-HIV, antiallergic,
CC antirheumatic, virucide, antiarthritic, ophthalmological and antimicrobial
CC activity. RNAAP antibodies are useful for diagnosis of diseases
CC associated with altered expression or activity of RNAAP. It is used to
CC treat cell proliferative, autoimmune, inflammatory and infectious
CC disorders, like actinic keratosis, bursitis, arteriosclerosis,
CC artherosclerosis, cirrhosis, hepatitis, myelofibrosis, mixed connective
CC tissue disease (MCTD), psoriasis, primary thrombocythemia and cancer,
CC HIV, allergies, rheumatoid arthritis, uveitis, Crohn's disease, and
CC bacterial, viral and parasitic infections.

XX Sequence 153 AA;

Query Match 74.5%; Score 38; DB 21; Length 153;

Best Local Similarity 50.0%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13

Db 141 ccgnsycdec 150

RESULT 12

ABB69415

ID ABB69415 standard; Protein; 237 AA.

XX AC ABB69415;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 35037.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX

PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR N-PSDB; ABL13518.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 35037; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 237 AA;
SQ

Query Match 74.5%; Score 38; DB 22; Length 237;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
Db 165 ccgkqncgtc 174
||| | |

RESULT 13
AAM41045
ID AAM41045 standard; Protein; 973 AA.
XX
AC AAM41045;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 5976.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR

PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI60201.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX
PS Example 2; SEQ ID NO 5976; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 973 AA;
SQ

Query Match 74.5%; Score 38; DB 22; Length 973;
Best Local Similarity 50.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
Db 122 ccgnsycdec 131
||| | |

RESULT 14
AAB11436
ID AAB11436 standard; Protein; 1023 AA.
XX
AC AAB11436;
XX
XX 23-FEB-2001 (first entry)
DT
XX D. discoideum sterol glucosyltransferase protein.
DE
XX Sterol glucosyltransferase; transgenic; sterol glycoside; cardioactive;
KW sterol alkaloid glycoside; hypocholesterolemic; hemolytic; fungicidal;
KW insecticidal; treatment; heart disease; saponin; disogenic; additive;
KW steroid hormone; emulsifier; dispersant; wetting agent.
XX
XX Dictyostellum discoideum.
OS
XX DE19919550-A1.
XX
XX 02-NOV-2000.
PD
XX 29-APR-1999; 99DE-1019550.
XX
XX 29-APR-1999; 99DE-1019550.
XX

PA (NDEP-) NORDEUT PFLANZENZUCHT LEMBKE KG HANS.

XX Heinz E, Warnecke D, Zank T, Mueller F, Zaehring U;

XX WPI: 2001-000472/01.

XX N-PSDB; AAC81909.

XX Preparation of transgenic organisms expressing sterol
PT glycosyltransferase, useful e.g. for high-yield production of
PT cardioactive steroid glycosides -

XX Examples: Page 44-47; 48pp; German.

XX This invention describes a novel preparation of transgenic organisms and
CC cells (A) with increased production of sterol glycosides (I), sterol
CC alkaloid glycosides (II) and/or sterol glycosides (III). The invention
CC also describes the construction of a nucleic acid sequence (IV) which
CC contains, in the 5' to 3' direction, a promoter (constitutive,
CC tissue-specific, development stage-specific or inducible) that is
CC functional in the appropriate host cell, at least 1 sequence (V) encoding
CC a sterol glycosyltransferase (SGT), preferably a sterol
CC glycosyltransferase, or its active fragment, and signals that terminate
CC transcription and add a polyA tail to the transcript or derived DNA. (IV)
CC is then introduced into a bacterial, fungal, plant or animal cell (and
CC optionally integrated into the genome) and optionally the cells are
CC regenerated into a complete transgenic organism which may be replicated.
CC The products of the invention have cardioactive, hypocholesterolemic,
CC hemolytic, fungicidal and insecticidal activity. (A) are used for
CC producing (I)-(III), (I) is used for treating heart disease, saponins are
CC for use as hypocholesterolemic and/or hemolytic agents, disogenin for
CC synthesis of sterol hormones (cortisone or progesterone). (I)-(III) are
CC also useful as emulsifiers, additives for washing compositions,
CC dispersants and wetting agents, to generate plants resistant to insects
CC and fungi, to produce Saccharomyces strains for baking, to increase the
CC glycolipid content of the lecithin fraction of vegetable oils, and to
CC produce transgenic animals that produce milk and eggs of reduced
CC cholesterol content. Nucleic acid that encodes sterol glycosyltransferase
CC (SGT) is useful for identifying specific effectors (particularly
CC inhibitors of the fungal enzyme and potential fungicides) and for in
CC vitro glycosylation of steroid hormones (to produce a slow release
CC pharmaceutical). (A) produce greater yields of (I)-(III) than wild-type
CC organisms or cells.

XX Sequence 1023 AA;

Query Match 74.5%; Score 38; DB 22; Length 1023;
Best Local Similarity 50.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13

||| |

Db 975 ccgkifesc 984

RESULT 15

AAM39259

ID AAM39259 standard; Protein; 1278 AA.

XX AC AAM39259;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2404.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX Homo sapiens.

XX

PN WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB; AAI58415.

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX Example 4; SEQ ID NO 2404; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX Sequence 1278 AA;

Query Match 74.5%; Score 38; DB 22; Length 1278;

Best Local Similarity 50.0%; Pred. No. 6e+02;

Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13

||| |

Db 98 ccgnsycdec 107

Search completed: May 19, 2002, 04:43:16
Job time: 12543 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 04:43:19 ; Search time 112.09 Seconds
(without alignments)
21.607 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCCGXXXXXXCX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phase:**
- 10: sp_plant:**
- 11: sp_rhodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_rvirus:**
- 16: sp_bacteriap:**
- 17: sp_archheap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	78.4	460	11 Q63175	Q63175 rattus norv
2	39	76.5	170	6 Q9TT27	Q9TT27 bos taurus
3	39	76.5	170	6 Q9TT26	Q9TT26 ovis aries
4	39	76.5	188	2 Q9TF29	Q9TF29 methylophag
5	39	76.5	246	4 Q99750	Q99750 homo sapien
6	39	76.5	246	11 Q99JM9	Q99JM9 mus musculu
7	39	76.5	286	3 Q06651	Q06651 saccharomyc
8	39	76.5	350	5 Q9N8U4	Q9N8U4 trypanosoma
9	39	76.5	721	12 Q91LL5	Q91LL5 white spot
10	39	76.5	1190	5 Q62Z35	Q62Z35 caenorhabdi
11	39	76.5	1231	5 Q9XZ21	Q9XZ21 drosophila
12	38	74.5	237	5 Q9V9I3	Q9V9I3 drosophila
13	38	74.5	258	12 Q36453	Q36453 lymantria d
14	38	74.5	326	2 Q9WX86	Q9WX86 clostridium
15	38	74.5	328	16 Q97H26	Q97H26 clostridium
16	38	74.5	363	5 Q9N5Y3	Q9N5Y3 caenorhabdi

17	38	74.5	948	4 Q15290	Q15290 homo sapien
18	38	74.5	1023	5 Q9XYD4	Q9XYD4 dictyosteli
19	38	74.5	1591	11 P97868	P97868 mus musculu
20	38	74.5	1616	4 Q96PH3	Q96PH3 homo sapien
21	37	72.5	151	12 Q10614	Q10614 helicoverpa
22	37	72.5	167	11 Q9D122	Q9D122 mus musculu
23	37	72.5	168	11 Q9D732	Q9D732 mus musculu
24	37	72.5	186	11 Q64526	Q64526 mus musculu
25	37	72.5	191	11 Q9D3H7	Q9D3H7 mus musculu
26	37	72.5	195	11 Q9D141	Q9D141 mus musculu
27	37	72.5	244	12 Q9IBQ9	Q9IBQ9 spodoptera
28	37	72.5	264	16 Q97IX6	Q97IX6 clostridium
29	37	72.5	285	12 Q99H38	Q99H38 helicoverp
30	37	72.5	289	12 Q91BK7	Q91BK7 spodoptera
31	37	72.5	312	5 Q45799	Q45799 caenorhabdi
32	37	72.5	446	4 Q96BQ3	Q96BQ3 homo sapien
33	37	72.5	550	10 Q23187	Q23187 arabidopsis
34	37	72.5	693	6 Q27960	Q27960 bos taurus
35	36	70.6	64	5 Q9BPE9	Q9BPE9 conus penna
36	36	70.6	102	11 Q9D7Q2	Q9D7Q2 mus musculu
37	36	70.6	113	11 Q9D228	Q9D228 mus musculu
38	36	70.6	122	11 Q9D230	Q9D230 mus musculu
39	36	70.6	138	11 Q9D3H6	Q9D3H6 mus musculu
40	36	70.6	404	10 Q64875	Q64875 arabidopsis
41	36	70.6	514	10 Q9LD86	Q9LD86 arabidopsis
42	36	70.6	756	10 Q9M999	Q9M999 arabidopsis
43	35	68.6	127	5 Q9GYP2	Q9GYP2 caenorhabdi
44	35	68.6	428	10 Q9M2P5	Q9M2P5 arabidopsis
45	34	66.7	215	5 Q21162	Q21162 caenorhabdi

ALIGNMENTS

RESULT 1					
Q63175	ID	Q63175	PRELIMINARY;	PRT;	460 AA.
AC	Q63175;				
DT	01-NOV-1996 (Tremblrel. 01, Created)				
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)				
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)				
DE	DEVELOPMENTALLY REGULATED PROTEIN.				
GN	TPOL.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SPRAGUE-DAWLEY; TISSUE-TELENCEPHALON;				
RX	MEDLINE=97465542; PubMed=9326262;				
RA	Krueger W.H., Gonye G.E., Madison D.L., Murray K.E., Kumar M.,				
RA	Spoerel N., Pfeiffer S.E.;				
RT	"TPOL, a member of a novel protein family, is developmentally				
RT	regulated in cultured oligodendrocytes.;"				
RL	J. Neurochem. 69:1343-1355(1997).				
CC	- - COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).				
DR	EMBL; L20319; AAA41097.1; -				
DR	InterPro: IPR000564; 2Fe2S.ferredoxin.				
DR	PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.				
KW	Iron-sulfur.				
SQ	SEQUENCE 460 AA; 51855 MW; CB9DA60F66282FB3 CRC64;				

Query Match 78.4%; Score 40; DB 11; Length 460;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 CCGXXXXXXC 13
DB 13 CCGSAGCALC 22

```

RESULT 2
Q9TTZ7 ID Q9TTZ7 PRELIMINARY; PRT; 170 AA.
AC Q9TTZ7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE INHIBITOR OF MYOD FAMILY-FORM A (FRAGMENT).
GN I-MF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorimer J.E., Thomas M., Kirk S., Bass J.;
RT "I-mfa, an inhibitor of the Myod family, is differentially expressed
during development.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184162; AAF05470.1; -.
DR InterPro; IPR001368; TNFR_c6.
DR InterPro; IPR001007; VFWC.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01208; VFWC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 170
SQ SEQUENCE. 170 AA; 17438 MW; 12695DCBB3060E2D CRC64;

Query Match 76.5%; Score 39; DB 6; Length 170;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXXC 13
DB 130 CCGGECADC 139

RESULT 3
Q9TTZ6 ID Q9TTZ6 PRELIMINARY; PRT; 170 AA.
AC Q9TTZ6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE INHIBITOR OF MYOD FAMILY-FORM A (FRAGMENT).
GN I-MF.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorimer J.E., Thomas M., Kirk S., Bass J.;
RT "I-mfa, an inhibitor of the Myod family, is differentially expressed
during development.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184163; AAF05471.1; -.
DR InterPro; IPR001368; TNFR_c6.
DR InterPro; IPR001007; VFWC.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01208; VFWC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 170
SQ SEQUENCE. 170 AA; 17507 MW; 99C048C534BBE79F CRC64;

Query Match 76.5%; Score 39; DB 6; Length 170;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXXC 13
DB 130 CCGGECADC 139

Query Match 76.5%; Score 39; DB 2; Length 188;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXXC 13
DB 134 CCGKQTCGSC 143

RESULT 5
Q99750 ID Q99750 PRELIMINARY; PRT; 246 AA.
AC Q99750;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOGENIC REPRESSOR I-MF (MYOD FAMILY INHIBITOR).
GN MDFI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=96390847; PubMed=8797820;
RA Chen C.M., Kraut N., Groudine M., Weintraub H.;
RT "I-mf, a novel myogenic repressor, interacts with members of the Myod
family.";
RL Cell 86:731-741(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE, RHABDOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78313; AAB39748.1; -.
DR EMBL; AL035588; CAB54148.1; -.
DR EMBL; BC007836; AAH07836.1; -.
DR InterPro; IPR001368; TNFR_c6.
DR InterPro; IPR001007; VFWC.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01208; VFWC; UNKNOWN_1.

```

```

QY 4 CCGXXCXXC 13
DB 130 CCGGECADC 139

RESULT 4
Q93FZ9 ID Q93FZ9 PRELIMINARY; PRT; 188 AA.
AC Q93FZ9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE METHYLAMINE DEHYDROGENASE SMALL SUBUNIT.
OS Methylophaga thalassica.
OC Bacteria; Proteobacteria; gamma subdivision; Piscirickettsia group;
OC Methylophaga.
OX NCBI_TaxID=40223;
RN [1]
RP SEQUENCE FROM N.A.
RA Chistoserdov A.Y.;
RT "Biochemical and genetic characterization of methylamine dehydrogenase
from the marine methylophag, Methylophaga thalassica sl.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF302651; AAL09314.1; -.
SQ SEQUENCE. 188 AA; 20301 MW; A8167654402AA428 CRC64;

Query Match 76.5%; Score 39; DB 2; Length 188;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXXC 13
DB 134 CCGKQTCGSC 143

RESULT 5
Q99750 ID Q99750 PRELIMINARY; PRT; 246 AA.
AC Q99750;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOGENIC REPRESSOR I-MF (MYOD FAMILY INHIBITOR).
GN MDFI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=96390847; PubMed=8797820;
RA Chen C.M., Kraut N., Groudine M., Weintraub H.;
RT "I-mf, a novel myogenic repressor, interacts with members of the Myod
family.";
RL Cell 86:731-741(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE, RHABDOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78313; AAB39748.1; -.
DR EMBL; AL035588; CAB54148.1; -.
DR EMBL; BC007836; AAH07836.1; -.
DR InterPro; IPR001368; TNFR_c6.
DR InterPro; IPR001007; VFWC.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01208; VFWC; UNKNOWN_1.

```

SQ SEQUENCE 246 AA; 25029 MW; 9330E52141B79E5A CRC64;

Query Match 76.5%; Score 39; DB 4; Length 246;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
| | | | |

Db 203 CCGSGECADC 212

RESULT 6

Q99JM9 ID Q99JM9 PRELIMINARY; PRT; 246 AA.

AC Q99JM9; DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOD FAMILY INHIBITOR.

GN MDFI.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-BREAST TUMOR;

RA Strausberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC006018; AAH06018.1; -

DR EMBL; BC010259; AAH10259.1; -

DR MGD; MGI:107687; Mdf1.

DR InterPro; IPR001368; TNFR_c6.

DR InterPro; IPR001007; VWFC.

DR PROSITE; PS00652; TNFR_NGFR.1; UNKNOWN_1.

DR PROSITE; PS01208; VWFC; UNKNOWN_1.

SQ SEQUENCE 246 AA; 25270 MW; 138A37A2BA4A10463 CRC64;

Query Match 76.5%; Score 39; DB 11; Length 246;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
| | | | |

Db 203 CCGSGECADC 212

RESULT 7

Q06651 ID Q06651 PRELIMINARY; PRT; 286 AA.

AC Q06651; DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE SIMILAR TO PROTEINS INVOLVED IN VACUOLAR FUNCTION: S. CEREVISIAE

DE VAC1P.

GN PIB1 OR D9740.20 OR YDR313C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,

RA Favetto A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,

RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,

RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,

RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA Taich A., Trevasakis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
RA Wilson R., Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ding H.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Jia Y., Cherry J.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28374; AAB64749.1; -
DR SGD; S0002721; PIB1.
DR InterPro; IPR000306; Znf_FYVE.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF01363; FYVE; 1.
DR SMART; SM00064; FYVE; 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 286 AA; 32675 MW; E7E9DCDEC9BAA183 CRC64;

Query Match 76.5%; Score 39; DB 3; Length 286;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
| | | | |

Db 41 CCGRIFCSC 50

RESULT 8

Q9N8U4 ID Q9N8U4 PRELIMINARY; PRT; 350 AA.

AC Q9N8U4;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE HYPOTHETICAL 38.0 KDA PROTEIN.

GN CHRI.143.

OS Trypanosoma brucei.

OC Eukaryota; Euclenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

OX NCBI_TaxID=5691;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=TREU927;

RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,

RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,

RA Gerrard C., Rajandream M.A., Barrell B.G.;

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL359782; CAB95439.1; -

KW Hypothetical protein.

SQ SEQUENCE 350 AA; 37995 MW; 98796C546331D2A4 CRC64;

Query Match 76.5%; Score 39; DB 5; Length 350;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
| | | | |

Db 234 CCGAIMCGAC 243

RESULT 9

Q91LL5 ID Q91LL5 PRELIMINARY; PRT; 721 AA.

AC Q91LL5;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ORF24.
OS White spot syndrome virus (WSSV).
OC Viruses: Unassigned viruses.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342572; PubMed=11448154;
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF369029; AAK77693.1; -.
RL "The white spot syndrome virus DNA genome sequence.";
RL Virology 286:7-22(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF369029; AAK77693.1; -.
SQ SEQUENCE 721 AA; 81137 MW; B1F9850F3CEA494D CRC64;

Query Match 76.5%; Score 39; DB 12; Length 721;
Best Local Similarity 50.0%; Pred.No. 40;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
Db 212 CCGAKICSEC 221

RESULT 10
O62235 PRELIMINARY; PRT; 1190 AA.
AC O62235;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE F36F2.3 PROTEIN.
GN F36F2.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Cottage A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT Investigating biology.";
RT Science 282:2012-2018(1998).
DR EMBL: Z81533; CAB04326.2; -.
SQ SEQUENCE 1190 AA; 137239 MW; B219D269962432AB CRC64;

Query Match 76.5%; Score 39; DB 5; Length 1190;
Best Local Similarity 50.0%; Pred.No. 53;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
Db 323 CCGNSYCADC 332

RESULT 11
Q9XZ21 PRELIMINARY; PRT; 1231 AA.
ID Q9XZ21
AC Q9XZ21;
DT 01-NOV-1999 (TReMBLrel. 12, Created)

```

```

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE BCDNA:LD21643 OR CG3231.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Renington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Celniker S.E.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AE003463; AAF47162.1; -.
DR EMBL: AF132177; AAD34765.1; -.
DR FlyBase: FBgn0027522; BcdNA:LD21643.
DR InterPro: IPR001878; Znf.CCHC.
DR InterPro: IPR001841; Znf_fing.
DR Pfam: PF00097; zf-C3HC4; 1.
DR Pfam: PF00098; zf-CCHC; 1.
DR SMART: SM00184; RING; 1.
DR SMART: SM00343; Znf_C2HC; 1.
KW Hypothetical protein; Zinc-finger.
SQ SEQUENCE 1231 AA; 139038 MW; 4A991FF172BDE4B0 CRC64;

```

Query Match 76.5%; Score 39; DB 5; Length 1231;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| |
Db 232 CCGSFCDDC 241

RESULT 12

Q9V9I3 PRELIMINARY; PRT; 237 AA.
AC Q9V9I3;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE CGL4468 PROTEIN.
GN TSP42A OR CGL4468.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003785; AAF57306.1; -;
DR FlyBase; FBgn0033042; Tsp42A.
DR InterPro; IPR000301; Transmen_4.
SQ SEQUENCE 237 AA; 27034 MW; 43E964F0E721D962 CRC64;

Query Match 74.5%; Score 38; DB 5; Length 237;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| |
Db 165 CCGKQNCCTGC 174

RESULT 13

O36453 PRELIMINARY; PRT; 258 AA.
ID O36453; Q9YMV3;
AC O36453; Q9YMV3;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE IMMEDIATE EARLY 0 PROTEIN (IMMEDIATE EARLY TRANSACTIVATOR 0).
GN IE-0.
OS Lymantria dispar multicapsid nuclear polyhedrosis virus (LdMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10449;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445058; PubMed=9300047;
RA Pearson M.N., Rohrmann G.F.;
RT "Splicing is required for transactivation by the immediate early gene
1 of the Lymantria dispar multinucleocapsid nuclear polyhedrosis
virus";
RL Virology 235:153-165(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99124785; PubMed=9887315;
RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavicek J.M., Rohrmann G.F.;
RT "Sequence and analysis of the genome of a baculovirus pathogenic for
Lymantria dispar";
RL Virology 253:17-34(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
RA Slavicek J., Rohrmann G.F.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Kuzio J.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006656; AAC58234.1; -;
DR EMBL; AF081810; AAC70206.1; -;
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 258 AA; 29395 MW; AC564CDF92282BAD CRC64;

Query Match 74.5%; Score 38; DB 12; Length 258;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCCXC 13
||| |
Db 209 CCGYRVCNAC 218

RESULT 14

Q9WX86 PRELIMINARY; PRT; 326 AA.
ID Q9WX86;
AC Q9WX86;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE ELECTRON TRANSFER SUBUNIT PROTEIN.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

Search completed: May 19, 2002, 08:06:54
Job time: 12215 sec

```
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PB6K;
RX MEDLINE=20090462; PubMed=10627036;
RA Fujinaga K., Taniguchi Y., Sun Y., Katayama S., Minami J.,
RA Matsushita O., Okabe A.;
RT "Analysis of genes involved in nitrate reduction in Clostridium
RT perfringens.";
RL Microbiology 145:3377-3387(1999).
CC -1- COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).
DR EMBL; AB017192; BAA74788.1; -.
DR InterPro; IPR000564; 2Fe2S-ferredoxin.
DR InterPro; IPR001281; Rieske.
DR PROSITE; PS00197; 2FE2S-FERREDOXIN; 1.
DR PROSITE; PS00200; RIESKE_2; UNKNOWN_1.
KW Iron-sulfur
SQ SEQUENCE 326 AA; 36522 MW; 2BDC1798F376FD65 CRC64;

Query Match 74.5%; Score 38; DB 2; Length 326;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXC 13
Db 288 CCGEGVCGSC 297

RESULT 15
Q97H26
ID Q97H26 PRELIMINARY; PRT; 328 AA.
AC Q97H26;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PREDICTED FLAVODOXIN OXIDOREDUCTASE.
GN CAC1860.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007694; AAK79824.1; -.
DR InterPro; IPR000564; 2Fe2S-ferredoxin.
DR InterPro; IPR001281; Rieske.
DR PROSITE; PS00197; 2FE2S-FERREDOXIN; 1.
DR PROSITE; PS00200; RIESKE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 328 AA; 37270 MW; 2417A9CEC7058AA1 CRC64;

Query Match 74.5%; Score 38; DB 16; Length 328;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXC 13
Db 290 CCGEGVCGAC 299
```


THIS PAGE BLANK (USPTO)

Query Match 80.4%; Score 41; DB 6; Length 2014;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 4 CCGXXXCXXC 13
Db 54 ccgtatctgc 63

Search completed: May 19, 2002, 08:04:48
Job time: 15749 sec

```
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 181
; LENGTH: 1624
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1560-1561
; OTHER INFORMATION: unknown base
US-10-123-155-181
```

```
Query Match 80.4%; Score 41; DB 6; Length 1624;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 4 CCGXXXCXXC 13
||| |
Db 109 ccgagtcac 118
```

```
RESULT 13
US-10-063-545-53
; Sequence 53, Application US/10063545
; GENERAL INFORMATION:
; APPLICANT: Eaton,Dan L.
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,Christopher J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230RIC1
; CURRENT APPLICATION NUMBER: US/10/063,545
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 53
; LENGTH: 1676
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-545-53
```

```
Query Match 80.4%; Score 41; DB 6; Length 1676;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 4 CCGXXXCXXC 13
||| |
Db 1274 ccgagtcac 1283
```

```
RESULT 14
US-10-063-685-53
; Sequence 53, Application US/10063685
; GENERAL INFORMATION:
; APPLICANT: Eaton,Dan L.
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,Christopher J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230RIC1
; CURRENT APPLICATION NUMBER: US/10/063,685
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 53
; LENGTH: 1676
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-685-53
```

```
Query Match 80.4%; Score 41; DB 6; Length 1676;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 4 CCGXXXCXXC 13
||| |
Db 1274 ccgagtcac 1283
```

```
RESULT 15
US-10-123-155-51
; Sequence 51, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Beresini,Maureen
; APPLICANT: DeForge,Laura
; APPLICANT: Desnoyers,Luc
; APPLICANT: Filvaroff,Ellen
; APPLICANT: Gao,Wei-Qiang
; APPLICANT: Gerritsen,Mary E.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Sherwood,Steven
; APPLICANT: Smith,Victoria
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Tumas,Daniel
; APPLICANT: Watanabe,Colin K
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 51
; LENGTH: 2014
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-51
```

Query Match 80.4%; Score 41; DB 6; Length 1346;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXC 13
| | | | |
Db 336 ccggtacac 345

RESULT 9
US-10-123-155-3
; Sequence 3, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 3
; LENGTH: 1395
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-3

Query Match 80.4%; Score 41; DB 6; Length 1395;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXC 13
| | | | |
Db 160 ccggtctac 169

RESULT 10
US-10-121-062-581
; Sequence 581, Application US/10121062
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1

; CURRENT APPLICATION NUMBER: US/10/121,062
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 581
; LENGTH: 1435
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-062-581

Query Match 80.4%; Score 41; DB 6; Length 1435;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXC 13
| | | | |
Db 899 CCGTTTCGAC 908

RESULT 11
US-10-123-155-17
; Sequence 17, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 17
; LENGTH: 1528
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-17

Query Match 80.4%; Score 41; DB 6; Length 1528;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXC 13
| | | | |
Db 710 ccgtgactac 719

RESULT 12
US-10-123-155-181
; Sequence 181, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen

```
; APPLICANT: Wood,William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 77
; LENGTH: 3567
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-77

Query Match      82.4%; Score 42; DB 6; Length 3567;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 CCGXXCXXC 13
      ||| | |
Db 2670 ccggtcttc 2679

RESULT 6
US-10-123-155-441
; Sequence 441, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 441
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-441

Query Match      80.4%; Score 41; DB 6; Length 654;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 CCGXXCXXC 13
      ||| | |
Db 397 ccggttcaac 406

RESULT 7
US-10-123-155-173
; Sequence 173, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 481
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-481

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 173
; LENGTH: 1210
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-173

Query Match      80.4%; Score 41; DB 6; Length 1210;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 CCGXXCXXC 13
      ||| | |
Db 911 ccgaatcgac 920

RESULT 8
US-10-123-155-481
; Sequence 481, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 481
; LENGTH: 1346
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-481
```

APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330RIC30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 61
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-61

Query Match 82.4%; Score 42; DB 6; Length 2119;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
Db 318 ccgtaacatc 327

RESULT 3
US-10-123-155-151
Sequence 151, Application US/10123155
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330RIC30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 151
LENGTH: 2164
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-151

^ . ^

Query Match 82.4%; Score 42; DB 6; Length 2164;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
Db 713 ccgttactac 722

RESULT 4
US-10-123-155-79
Sequence 79, Application US/10123155
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330RIC30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 79
LENGTH: 2714
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-79

Query Match 82.4%; Score 42; DB 6; Length 2714;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
||| | |
Db 646 ccgatacaac 655

RESULT 5
US-10-123-155-77
Sequence 77, Application US/10123155
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2002, 03:42:19 ; Search time 37.95 Seconds
(without alignments)
9.751 Million cell updates/sec

Title: US-09-580-201A-1
Perfect score: 51
Sequence: 1 XXXCGXXXCXXCX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 107366 seqs, 26430961 residues

Total number of hits satisfying chosen parameters: 107366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	42	82.4	1041	6 US-10-123-155-21	Sequence 21, Appl
2	42	82.4	2119	6 US-10-123-155-61	Sequence 61, Appl
3	42	82.4	2164	6 US-10-123-155-151	Sequence 151, Appl
4	42	82.4	2714	6 US-10-123-155-79	Sequence 79, Appl
5	42	82.4	3567	6 US-10-123-155-77	Sequence 77, Appl
6	41	80.4	654	6 US-10-123-155-441	Sequence 441, Appl
7	41	80.4	1210	6 US-10-123-155-173	Sequence 173, Appl
8	41	80.4	1346	6 US-10-123-155-481	Sequence 481, Appl
9	41	80.4	1395	6 US-10-123-155-3	Sequence 3, Appl
10	41	80.4	1435	6 US-10-121-062-581	Sequence 581, Appl
11	41	80.4	1528	6 US-10-123-155-17	Sequence 17, Appl
12	41	80.4	1624	6 US-10-123-155-181	Sequence 181, Appl
13	41	80.4	1676	6 US-10-063-545-53	Sequence 53, Appl
14	41	80.4	1676	6 US-10-063-685-53	Sequence 53, Appl
15	41	80.4	2014	6 US-10-123-155-51	Sequence 51, Appl
16	41	80.4	2061	6 US-10-063-545-131	Sequence 131, Appl
17	41	80.4	2061	6 US-10-063-685-131	Sequence 131, Appl
18	41	80.4	2236	6 US-10-123-155-381	Sequence 381, Appl
19	41	80.4	2285	6 US-10-123-155-93	Sequence 93, Appl
20	41	80.4	2379	6 US-10-063-545-535	Sequence 535, Appl
21	41	80.4	2379	6 US-10-063-545-123	Sequence 123, Appl
22	41	80.4	2379	6 US-10-063-685-123	Sequence 123, Appl
23	41	80.4	2668	6 US-10-123-155-511	Sequence 511, Appl
24	41	80.4	2846	6 US-10-063-545-37	Sequence 37, Appl
25	41	80.4	2846	6 US-10-063-685-37	Sequence 37, Appl
26	41	80.4	2848	6 US-10-063-545-97	Sequence 97, Appl

27	41	80.4	2848	6 US-10-063-685-97	Sequence 97, Appl
28	41	80.4	3334	6 US-10-063-545-57	Sequence 57, Appl
29	41	80.4	3334	6 US-10-063-685-57	Sequence 57, Appl
30	41	80.4	3819	6 US-10-123-155-405	Sequence 405, Appl
31	41	80.4	3884	6 US-10-123-155-145	Sequence 145, Appl
32	41	80.4	4374	6 US-10-123-155-125	Sequence 125, Appl
33	41	80.4	4420	6 US-10-123-155-411	Sequence 411, Appl
34	40	78.4	210	6 US-10-114-893-107	Sequence 107, Appl
35	40	78.4	422	6 US-10-063-545-65	Sequence 65, Appl
36	40	78.4	422	6 US-10-063-685-65	Sequence 65, Appl
37	40	78.4	526	6 US-10-123-155-249	Sequence 249, Appl
38	40	78.4	754	6 US-10-063-545-141	Sequence 141, Appl
39	40	78.4	754	6 US-10-063-685-141	Sequence 141, Appl
40	40	78.4	823	6 US-10-123-155-379	Sequence 379, Appl
41	40	78.4	1024	6 US-10-123-155-315	Sequence 315, Appl
42	40	78.4	1234	6 US-10-063-545-63	Sequence 63, Appl
43	40	78.4	1234	6 US-10-063-685-63	Sequence 63, Appl
44	40	78.4	1281	6 US-10-123-155-471	Sequence 471, Appl
45	40	78.4	1395	6 US-10-063-545-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-10-123-155-21
; Sequence 21, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 21
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-155-21

Query Match 82.4%; Score 42; DB 6; Length 1041;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXXC 13
DB 331 ccgtatcatc 340

RESULT 2
US-10-123-155-61
; Sequence 61, Application US/10123155
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

; APPLICANT: The Procter & Gamble Company
; APPLICANT: Smerznak, Mark A.
; APPLICANT: Broeckx, Walter A. M.
; APPLICANT: Johnston, James P.
; APPLICANT: Fredj, Abdennaceur
; TITLE OF INVENTION: Low Density Enzyme Granulates and Compositions Employing Same
; FILE REFERENCE: 7410
; CURRENT APPLICATION NUMBER: US/09/890,677
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1497
; TYPE: PRT
; ORGANISM: bacillus amyloliquefaciens
US-09-890-677-3

Query Match 80.4%; Score 41; DB 22; Length 1497;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXC 13
||| |
Db 841 CCGTTGCATC 850

RESULT 15
PCT-US01-12836-26
; Sequence 26, Application PC/TUS0112836
; GENERAL INFORMATION:
; APPLICANT: Egrie, Joan
; APPLICANT: Elliott, Steven
; APPLICANT: Browne, Jeffrey
; APPLICANT: Karen, Sitney
; TITLE OF INVENTION: Methods and Compositions for the Prevention and Treatment of Aneurysm
; FILE REFERENCE: A-460A
; CURRENT APPLICATION NUMBER: PCT/US01/12836
; CURRENT FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 09/178,292
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 2572
; TYPE: PRT
; ORGANISM: Human
PCT-US01-12836-26

Query Match 80.4%; Score 41; DB 1; Length 2572;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXXCXC 13
||| |
Db 1003 CCGAGTCTAC 1012

Search completed: May 19, 2002, 08:00:18
Job time: 21700 sec

```

; APPLICANT: Estell, David Aaron
; TITLE OF INVENTION: Human Protease and Use of Such Protease for Pharmaceutical
; TITLE OF INVENTION: Applications and for Reducing the Allergenicity of Non-Human
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: GC532
; CURRENT FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1497
; TYPE: PRT
; ORGANISM: B. amyloliquefaciens
US-09-060-854-2

```

```

Query Match      80.4%; Score 41; DB 14; Length 1497;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 CCGXXXCCXC 13
    ||| | |
Db 841 CCGTTCATC 850

```

```

RESULT 11
US-09-060-854B-2
; Sequence 2, Application US/09060854B
; GENERAL INFORMATION:
; APPLICANT: Estell, David Aaron
; TITLE OF INVENTION: Human Protease and Use of Such Protease for Pharmaceutical
; TITLE OF INVENTION: Applications and for Reducing the Allergenicity of Non-Human
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: GC532
; CURRENT FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1497
; TYPE: PRT
; ORGANISM: B. amyloliquefaciens
US-09-060-854B-2

```

```

Query Match      80.4%; Score 41; DB 14; Length 1497;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 CCGXXXCCXC 13
    ||| | |
Db 841 CCGTTCATC 850

```

```

RESULT 12
US-09-178-173-2
; Sequence 2, Application US/09178173
; GENERAL INFORMATION:
; APPLICANT: Poulos, Ayrookaran J.
; APPLICANT: Schellenberger, Volker
; APPLICANT: Kellis, Jr., James T.
; APPLICANT: Paech, Christian
; APPLICANT: Naki, Donald P.
; APPLICANT: Collier, Katherine D.
; APPLICANT: Caldwell, Robert M.

```

```

; APPLICANT: Baek, Andre C.
; TITLE OF INVENTION: Multiply-Substituted Protease Variants
; FILE REFERENCE: GC503-2
; CURRENT APPLICATION NUMBER: US/09/178,173
; CURRENT FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1497
; TYPE: PRT
; ORGANISM: B. amyloliquefaciens
US-09-178-173-2

```

```

Query Match      80.4%; Score 41; DB 15; Length 1497;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 CCGXXXCCXC 13
    ||| | |
Db 841 CCGTTCATC 850

```

```

RESULT 13
US-09-529-904-3
; Sequence 3, Application US/09529904
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Chanchal K.
; APPLICANT: Baek, Andre C.
; APPLICANT: Ohtani, Ryohei (nmn)
; APPLICANT: Busch, Alfred (nmn)
; APPLICANT: Showell, Michael S.
; APPLICANT: Poulos, Ayrookaran J.
; APPLICANT: Schellenberger, Volker (nmn)
; APPLICANT: Kellis, Jr., James T.
; APPLICANT: Paech, Christian (nmn)
; APPLICANT: Nadherny, Joanne (nmn)
; APPLICANT: Naki, Donald P.
; APPLICANT: Collier, Katherine D.
; APPLICANT: Caldwell, Robert M.
; TITLE OF INVENTION: Bleaching compositions comprising multiply-substituted
; TITLE OF INVENTION: protease variants
; FILE REFERENCE: Bleaching comp comprising multiply-sub
; CURRENT APPLICATION NUMBER: US/09/529,904
; CURRENT FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 08/956,323
; PRIOR FILING DATE: 1997-10-23
; PRIOR APPLICATION NUMBER: US 08/956,564
; PRIOR FILING DATE: 1997-10-23
; PRIOR APPLICATION NUMBER: US 08/956,324
; PRIOR FILING DATE: 1997-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1497
; TYPE: PRT
; ORGANISM: Bacillus amyloliquefaciens
US-09-529-904-3

```

```

Query Match      80.4%; Score 41; DB 19; Length 1497;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 CCGXXXCCXC 13
    ||| | |
Db 841 CCGTTCATC 850

```

```

RESULT 14
US-09-890-677-3
; Sequence 3, Application US/09890677
; GENERAL INFORMATION:

```


;
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486

;
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087098
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087208
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088722
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18

Query Match 80.4%; Score 41; DB 24; Length 1435;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

; Sequence 1, Application US/09735933
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000863
; CURRENT APPLICATION NUMBER: US/09/735,933
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1422
; TYPE: PRT
; ORGANISM: Human
US-09-735-933-1

Query Match 82.4%; Score 42; DB 21; Length 1422;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXCXC 13
||| | |
Db 275 CCGTATCAAC 284

RESULT 7
US-10-073-912-17
; Sequence 17, Application US/10073912
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ34C1
; CURRENT APPLICATION NUMBER: US/10/073,912
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 18636
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-073-912-17

Query Match 82.4%; Score 42; DB 24; Length 18636;
Best Local Similarity 50.0%; Pred. No. 3.7e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXCXC 13
||| | |
Db 10155 ccgtttctac 10164

RESULT 8
US-60-187-409-38
; Sequence 38, Application US/60187409
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: US\$ THEREOF
; FILE REFERENCE: CL000332
; CURRENT APPLICATION NUMBER: US/60/187,409
; CURRENT FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 606
; TYPE: PRT
; ORGANISM: HUMAN
US-60-187-409-38

Query Match 80.4%; Score 41; DB 26; Length 606;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CCGXXCXCXC 13
||| | |
Db 276 CCGTGTCTTC 285

RESULT 9
US-10-052-586-581
; Sequence 581, Application US/10052586
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18

```
RESULT 2
US-09-500-746-20
; Sequence 20, Application US/09500746
; GENERAL INFORMATION:
; APPLICANT: Winchester , Robert J.
; APPLICANT: Gulko, Percio
; TITLE OF INVENTION: USES OF INHIBITORS FOR THE ACTIVATION OF CXCR4 RECEPTOR BY SDF-1
; FILE REFERENCE: 0575/57005-B
; CURRENT APPLICATION NUMBER: US/09/500,746
; CURRENT FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Human
US-09-500-746-20

Query Match      82.4%; Score 42; DB 19; Length 377;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13
Db 13 CCGATTCTTC 22

RESULT 3
US-09-500-746-18
; Sequence 18, Application US/09500746
; GENERAL INFORMATION:
; APPLICANT: Winchester , Robert J.
; APPLICANT: Gulko, Percio
; TITLE OF INVENTION: USES OF INHIBITORS FOR THE ACTIVATION OF CXCR4 RECEPTOR BY SDF-1
; FILE REFERENCE: 0575/57005-B
; CURRENT APPLICATION NUMBER: US/09/500,746
; CURRENT FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Human
US-09-500-746-18

Query Match      82.4%; Score 42; DB 19; Length 377;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13
Db 13 CCGATTCTTC 22

RESULT 4
US-09-500-746-23
; Sequence 23, Application US/09500746
; GENERAL INFORMATION:
; APPLICANT: Winchester , Robert J.
; APPLICANT: Gulko, Percio
; TITLE OF INVENTION: USES OF INHIBITORS FOR THE ACTIVATION OF CXCR4 RECEPTOR BY SDF-1
; FILE REFERENCE: 0575/57005-B
; CURRENT APPLICATION NUMBER: US/09/500,746
; CURRENT FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 23
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)..(2)
; OTHER INFORMATION: X = to any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (3)..(3)
; OTHER INFORMATION: X = to any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (20)..(20)
; OTHER INFORMATION: X = to any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (22)..(22)
; OTHER INFORMATION: X = to any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (69)..(69)
; OTHER INFORMATION: X = to any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (70)..(70)
; OTHER INFORMATION: X = to any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (78)..(78)
; OTHER INFORMATION: X = to any amino acid
; NAME/KEY: MISC_FEATURE
; LOCATION: (410)..(410)
; OTHER INFORMATION: X = to any amino acid
US-09-500-746-23

Query Match      82.4%; Score 42; DB 19; Length 462;
Best Local Similarity 50.0%; Pred. No. 6.9e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13
Db 93 CCGATTCTTC 102

RESULT 5
US-10-048-394-3
; Sequence 3, Application US/10048394
; GENERAL INFORMATION:
; APPLICANT: Joele Thonnard
; TITLE OF INVENTION: Novel Compounds
; FILE REFERENCE: BM45408
; CURRENT APPLICATION NUMBER: US/10/048,394
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 9918206.5
; PRIOR FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-048-394-3

Query Match      82.4%; Score 42; DB 24; Length 648;
Best Local Similarity 50.0%; Pred. No. 8.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CCGXXXCXXC 13
Db 266 CCGAATCTAC 275

RESULT 6
US-09-735-933-1
```

Sequence 17, Appl